Contributions to energy-efficient wireless communications exploiting games, online optimization and learning

Elena Veronica Belmega

Contributions to energy-efficient wireless communications exploiting games, online optimization and learning

Contributions aux communications sans fil efficaces en énergie exploitant la théorie des jeux, l'optimisation en ligne et l'apprentissage

By

E. VERONICA BELMEGA

Maître de Conférences à l'ENSEA

ETIS UMR 8051, Université Paris Seine, Université de Cergy-Pontoise, ENSEA, CNRS

HABILITATION À DIRIGER DES RECHERCHES
DE L'UNIVERSITÉ DE CERGY-PONTOISE
Section CNU 61, génie informatique, automatique et traitement du signal

Defended on the 29th of March 2019 in front of the jury composed of

Pierre DUHAMEL	DR, CNRS	President of the jury (*Président*)
Marilyne HÉLARD	PU, INSA-Rennes	Reviewer (*Rapportrice*)
Giuseppe CAIRE	Full Professor, TU Berlin	Reviewer (*Rapporteur*)
Jean-Marie GORCE	PU, INSA-Lyon	Reviewer (*Rapporteur*)
Geneviève BAUDOIN	PU, ESIEE Paris	Examiner (*Examinatrice*)
Sergio BARBAROSSA	Full Professor, Sapienza Univ. of Rome	Examiner (*Examinateur*)
Inbar FIJALKOW	PU, ENSEA	HDR Guarantor (*Garante HDR*)
Dan VODISLAV	PU, Univ. of Cergy-Pontoise	HDR Referent (*Référent HDR*)

Abstract

Energy efficiency has been identified as one of the major desiderata for 5G communications, and will likely remain a crucial aspect for the information and communication technology (ICT) industry beyond 5G, owing to the prolific spread of Internet-enabled mobiles coupled with connectivity- and data-hungry applications. Moreover, the emerging Internet of things (IoT) paradigm, envisioned to interconnect wireless "things" (wireless sensors, wearables, biochip transponders, etc.) at a massive scale, also comes with strict limitations in terms of computational capabilities and energy consumption.

The aim of this HDR thesis is to design efficient resource allocation policies (in terms of available power, spectrum, space, and/or time) in wireless networks of interconnected devices, potentially equipped with multiple antennas (MIMO) or operating over multiple orthogonal frequency subcarriers (OFDM). First, we focus on networks that are relatively static in time and investigate different energy efficiency metrics based on the tradeoff between power consumption and Shannon rate. We then propose different energy-efficient policies that exploit promising technologies to address the spectrum and energy gridlocks (e.g., cognitive radio, small cell networks, MIMO, energy harvesting) and rely on classic tools and solution concepts from convex optimization, in centralized networks, and non-cooperative games, in decision-wise distributed networks.

The second part of this work incorporates the temporal network dynamics. This aspect is particularly relevant in IoT networks, in which coping with the high temporal variability (potentially non-stationary) caused by the devices' heterogeneous characteristics (in terms of mobility, connectivity patterns, etc.) is a major challenge. Algorithms that target fixed operating points or network states (e.g., classic optimal solutions or Nash equilibria) are no longer relevant in such arbitrarily varying and unpredictable networks. Instead, drawing on new tools from online optimization and learning, we propose efficient and dynamic resource allocation policies that are able to adapt on-the-fly to the network changes, while relying only on strictly causal and limited feedback information.

The thesis is concluded by discussing several open and prospective research directions for future exploration.

RÉSUMÉ

L'efficacité énergétique a été identifiée comme l'un des objectifs majeurs pour les communications 5G, et restera probablement un aspect crucial de l'industrie des technologies de l'information et de la communication (ICT) au-delà de la 5G, en raison de la prolifération des applications Internet mobile très gourmandes en débit et connectivité. De plus, le nouveau paradigme de l'Internet des objets (IoT), qui prévoit d'interconnecter des "objets" sans fil (capteurs, dispositifs prêts-à-porter, biopuces connectées, etc.) de manière massive, a des limitations strictes en termes de capacités de calcul et de consommation d'énergie.

L'objectif de cette thèse d'HDR est de concevoir des politiques efficaces d'allocation de ressources (puissance, spectre, espace et/ou temps) dans des réseaux sans fils composés par des dispositifs équipés de multiples antennes (MIMO) ou communicant dans plusieurs sous-porteuses orthogonales (OFDM). Premièrement, nous nous concentrons sur les réseaux relativement statiques dans le temps afin d'étudier plusieurs métriques d'efficacité énergétique définies en fonction du compromis entre la consommation de puissance et le débit de Shannon. Nous proposons ensuite différentes politiques efficaces en énergie exploitant des technologies prometteuses pour combattre les limitations du spectre et d'énergie (par exemple, la radio cognitive, les réseaux à petites cellules, les réseaux MIMO, la récupération d'énergie sans fils) et reposant sur des outils et concepts classiques issus de l'optimisation convexe, dans des réseaux centralisés, et des jeux non-coopératifs, dans des réseaux distribués au sens de la décision.

La deuxième partie de ce manuscrit intègre la dynamique temporelle du réseau. Cet aspect est particulièrement pertinent dans les réseaux IoT, où la prise en compte de la grande variabilité temporelle (potentiellement non stationnaire) due aux caractéristiques hétérogènes des objets (en termes de mobilité, schémas de connection, etc.) constitue un défi majeur. Les algorithmes qui ciblent des points de fonctionnement ou des états fixes (par exemple, des solutions optimales classiques ou des équilibres de Nash) ne sont plus pertinents dans des réseaux aussi variables et imprévisibles. En nous appuyant sur de nouveaux outils d'optimisation et d'apprentissage en ligne, nous proposons des politiques d'allocation de ressources efficaces et dynamiques, capables de s'adapter à la volée aux changements du réseau, en se reposant uniquement sur des informations de retour strictement causales et limitées.

La thèse est conclue par une discussion sur plusieurs axes de recherche ouverts et prospectifs pour des investigations futures.

DEDICATION AND ACKNOWLEDGEMENTS

"Do not be afraid to ask yourself silly questions!"

Pierre Duhamel, 2019

I would like to thank the members of the jury – Geneviève Baudoin (ESIEE Paris), Sergio Barbarossa (Sapienza Univ. of Rome), Pierre Duhamel (L2S, CNRS), Inbar Fijalkow (ETIS, ENSEA) and Dan Vodislav (Univ. of Cergy-Pontoise) – and the three reviewers – Maryline Hélard (INSA-Rennes), Giuseppe Caire (TU Berlin), and Jean-Marie Gorce (INSA-Lyon) – for kindly accepting my invitation, complying with all the HDR admin constraints and honouring me with your participation, which lead to a very interesting and rich scientific discussion.

A big thanks is directed to all ETIS administrative officers: Annick, Astrid, Anthony, and Sokhena, to ENSEA's head of the research and transfer unit, Marine, and also to Anne from the research unit at the University of Cergy-Pontoise. Thank you for all your precious help and assistance, for often going beyond the job description and making our lives so much easier.

Being a part of ETIS and ENSEA since 2011 and of Inria (between 2015-2017) has been a great and fulfilling academic experience. I would like to thank my ICI team-mates (Iryna, Laura, Inbar, Maël, Ersi, Ligong, Marwa …) for your dynamism and enthusiasm, for being such a driving force! Inbar, I would like thank you so much for your amazing support, always encouraging me to push my limits, for your guidance in advising our students and for all the opportunities you have so generously provided to me from the start. My office-mates Aymeric and David, thank you for sharing the ups and downs, the laughters, the tears and the encouragements; my ETIS and ENSEA colleagues (Camille, Emmanuelle, Mathias, Myriam, Laurence, Antoine, Laurent, Thomas, Christophe, Hedi, Son …) you make this journey lighter and funnier every day.

I would also like to thank my PhD advisors (Samson and Mérouane), all my collaborators and co-authors (Vince, Walid, Panayotis, Lalitha, Giacomo, Ángeles, Are, Eitan, Yeze, Anne, Ersi, Subhash, Romain, Aris, Samir …) from such diverse backgrounds and horizons; our interactions shape me not only as a researcher but also as a person. In particular, I have been so lucky to work closely with Panayotis, Inbar, Mérouane, Luca and Ersi over the last years: thank you for everything that you have brought to me as a researcher and to my career, for our enriching, lively exchanges and scientific discussions – for our agreements and disagreements: I enjoy every moment and I hope our collaborations will continue as successfully! To my students: Raouia, Alex, Irched, Gada, Kimon, Camilla … thank you for your dedication, your hard work, and for never ceasing to surprise me!

All my kindest thoughts are also addressed to my Romanian and Marseillaise families and friends and my dearest friends Diarra, Consuela, and Anne: thank you for all your love and support.

Now, I would like to dedicate this HDR thesis to my "partner in crime" Romain: your unconditional love and support, your humour and creativity, your calmness, patience, unbreakable optimism and relentlessness have always been a precious source of inspiration, motivation and strength. Your help with various software (LaTeX) and technical issues has been brilliant throughout. Thank you for coping with me so bravely in the most difficult and the most happy moments; I am looking forward to new horizons and adventures together!

TABLE OF CONTENTS

List of Tables

List of Figures

Figure **Page**

ACRONYMS

5G 5th generation communications. 33

AI artificial intelligence. 85

AP access point. 48

ASE area spectral efficiency. 51

CSI channel state information. 46, 47, 67

DE Debreu equilibrium. xi, 47–52

DNN Deep neural networks. 88, 89

EE energy efficiency. xi, 34, 52, 68, 69, 71, 72

EH energy harvesting. 55–57

GNE generalized Nash equilibrium. 48

HetNet heterogeneous network. 34, 50

ICT information and communications technology. 33

IoT Internet of things. 34, 53, 73, 81–83

MAB multi-armed bandit. 62, 66, 67, 83, 86

MAI multiple access interference. 51

MIMO multiple-input and multiple-output. vii, 34, 67, 79, 83–85

MUE macrocell user equipment. xi, 51

NE Nash equilibrium. 55–57

NOMA non-orthogonal multiple access. 34, 82, 83

OFDM orthogonal frequency division multiplexing. 34, 57

OFDMA orthogonal frequency-division multiple access. 47

OGA online gradient ascent. xi, 71–73

OGD Online gradient descent. 64, 65, 70, 73

OMD Online mirror descent. 65, 66, 73, 75

OXL online exponential learning. xii, 76, 78

OXL$_0$ online exponential learning with zeroth-order feedback. xii, 78

QoS quality of service. 33, 47, 83

SC small-cell. 33, 34, 51, 52

SCA small-cell access point. 50

SE Stackelberg equilibrium. 55, 56

SINR signal-to-interference-plus-noise ratio. 48, 50

SIR signal-to-interference ratio. 55–57

SKG Secret key generation. 53, 56–58

SUE small-cell user equipment. xi, 51

UAV unmanned aerial vehicle. 81

UE user equipment. 33, 47–49, 51

WF water-filling. 50

Curriculum Vitae

This HDR thesis starts with a brief *Curriculum Vitae*, before delving into the detailed review of all my activities as a *Maître de Conférences* since September 2011. This *curriculum* offers a global and concise overview of my career path so far.

E. Veronica Belmega

Associate Professor (*Maître de Conférences*) at ENSEA

Address	ETIS/ENSEA, 6, Av. du Ponceau 95014 Cergy-Pontoise	**Office Phone**	+ 33 1 30 73 62 92
		Email	belmega@ensea.fr
Date of Birth	22^{nd} May 1983	**Webpage**	sites.google.com/site/evbelmega
Nationalities	Romanian and French	**Civil Status**	Married

Research Interests

Optimization, game theory and learning algorithms applied to MIMO systems, wireless communication, IoT networks, physical layer security, cognitive radio, and electrical power grids.

Professional Experience

Sep 2011 - Present Associate Professor (*Maître de Conferénces*) - ENSEA, Cergy-Pontoise, France
***Research**: ETIS, **Teaching**: Computer Science Dept. and Signal Processing Dept.*

Sep 2015 - Aug 2017 Visiting Researcher (Délégation complète) - Inria, Grenoble Rhône-Alpes, France
***Collaborator:** Panayotis Mertikopoulos*

Jan 2011 - Aug 2011 Post-doctoral Researcher - Princeton University (NJ, USA) and Supélec (Gif-sur-Yvette, France)
***Collaborators:** Vincent H. Poor, Lalitha Sankar, Walid Saad, Mérouane Debbah*

Sep 2007 - Aug 2010 Teaching Assistant (*Monitrice*) - École Polytechnique, Palaiseau, France
***Teaching:** Physics Dept.*

Education

2007-2010 Ph.D. in Telecommunications - Université Paris-Sud 11, Orsay, France
"On resource allocation problems in distributed MIMO wireless channels"
***Research:** L2S, **Advisors:** Samson Lasaulce, Mérouane Debbah*

2006-2007 M. Sc. in Signal Processing and Communications - Université Paris-Sud 11, Orsay, France
Rank: 1/13, *Mention TB - Trés Bien*

2005-2006 International Socrates/Erasmus Exchange Program - École Polytechniques, Palaiseau, France
***Track:** Applied Mathematics (3rd year, French engineer cycle)*

2002-2007 Engineer Diploma in Control and Computer Science - Politehnica University of Bucharest, Romania
***Rank:** 6/210*

Awards and Distinctions

2018 Doctoral Supervision and Research Bonus (PEDR) - French National Council of Universities (CNU 61)

2017 Top 11 Editors - Transactions on Emerging Telecommunications Technologies (ETT)

2015 Fellowship for Mobility (*Délégation*) - Inria, France

2012 Associate Research Scholar Fellowship - Princeton University, NJ, USA

2009 National Fellowship - L'Oréal France - UNESCO - French Academy of Science

2007 Ph.D. Fellowship - French Ministry of Research and Education

2005 Socrates/Erasmus International Scolarship - École Polytechnique and Politehnica University of Bucharest

Advising

PhDs 2 ongoing (75%, 30%) and 2 defended (40%, 50%)
Post-docs 2

Publications

21 (14*) International journal papers, **2 (1*)** invited
38 (27*) International conference papers, **11 (4*)** invited
4 (1*) Book chapters
5 (2*) National conference papers
* Publication of research work post-PhD

Scientific Responsibilities

2016 - 2018 Signal-Image-Vision Best PhD prize organized jointly by the French Club EEA, GdR ISIS and GRETSI
Jury member, Vice-president in 2017 and President in 2018

Jan. 2013 - Nov. 2017 GdR ISIS French Network in Image, Signal and Vision funded by the CNRS and French industrial partners
Scientific coordinator (Directrice scientifique adjointe)

Local Responsibilities

2013-2016 Member of the **Scientific Board** of ENSEA

2014, 2018 Co-chair of the Networks and Communications Track (*Option RT - Réseaux et Télécommunications*)

Editorial Activities

Jul. 2016 - present Transactions on Emerging Telecommunications Technologies (ETT)
Executive Editor

Feb. 2017 - Oct. 2017 IET Signal Processing
Associate Editor

Conference Organization

2017 GdR ISIS meeting: *"Game theory, optimization and learning"*, Paris, France

2014 International workshop: *"Wireless Networks, Communication, Cooperation and Competition" (WNC3)*, **WiOpt**, Hammamet, Tunisia

2013 Special session: "*Communications and Control in the Smart Grid" (CCSG)*, **IEEE BlackSeaCom**, Batumi, Georgia

CHAPTER 1

ACTIVITY REVIEW

Following the brief *Curriculum Vitae*, this chapter is dedicated to a detailed review of the activities I have been leading as a *Maître de Conférences* in ENSEA and ETIS lab, Cergy-Pontoise, France since September 2011. Both the teaching and the research activities will be overviewed, as well as the relevant administrative and scientific responsibilities.

1.1 Teaching Activities

My teaching activities at ENSEA are pluridisciplinary and lie at the intersection between two departments: the Computer Science Dept. and the Signal Processing Dept.. The specific department is automatically defined by the majority of the taught classes (each class being associated to one Dept.). Given the dynamics of my activities (illustrated in Table 1.3) and the wish to strengthen the ties between my teaching and my research activities, I have been associated first with the Computer Science Dept. (between Sep. 2011–Aug. 2015) and then with the Signal Processing Dept. (since Sep. 2017).

In between Sep. 2015 – Aug. 2017, I have been granted a mobility leave (*délégation à temps complet*) from Inria Grenoble Rhône-Alpes. As a result and given my mobility to Grenoble, France, I have been relieved from all teaching duties from ENSEA for two full years.

Tables 1.1, 1.2, 1.3 below offer a concise summary of my teaching activities in terms of: amount of hours (*équivalent TD*); student levels (all levels of the French engineering cycle and the local M2R); type (CM - *cours magistral*, TD - *travaux dirigés*, TP - *travaux pratiques et projets*); and taught subjects. The "Other" category represents the equivalent number of hours granted in exchange of various module and administrative responsibilities, student projects, etc.. A highlight point to be noted is the **high dynamics** of my activities in terms of new subjects per year.

Year	**CM**	**TD**	**TP**	**Other**	**Total**
2011 - 2012	12	6	194	8	220
2012 - 2013	60	23	132	7	222
2013 - 2014	81	23	130	18	252
2014 - 2015	77	22	56	51	206
2015 - 2016	-	-	-	-	-
2016 - 2017	-	-	-	-	-
2017 - 2018	45	34	108	37	224
2018 - 2019	36	46	92	24	198
Total	311	154	712	145	1322
Percentage	23.5%	11.6%	53.8%	11.1%	-

Table 1.1: Teaching at ENSEA during 2011–2019 in nb. hours (*équivalent TD*).

Subject	**Level**	**Type**	**Nb. hours**	**Percentage**
Advanced Digital Communications	M2	CM, TD, TP	152	11.5%
Introduction to Game Theory	M2	CM	14	1%
Communication Systems	M2	CM, TD	96	7.2%
Game Theory Applied to Communications	M1	CM, TP	90	6.8%
Introduction to Digital Communications	M1	CM, TD, TP	114	8.6%
Digital Signal Processing	M1	TD, TP	60	4.5%
Algorithms	M2	CM, TD, TP	219	16.5%
C++ Object Oriented Programming	M1	CM, TP	96	7.3%
JAVA Programming	M1	TP	84	6.3%
Data structures	L3	TP	72	5.4%
Computer science (C programming)	L3	TD, TP	156	11.8%
Microprocessors	L3	TP	24	1.8%
Other			145	11.1%

Table 1.2: Subjects taught at ENSEA during 2011–2019 in nb. hours (*équivalent TD*).

1.1.1 Proposed Modules

The three modules that **I have proposed (which did not exist) and prepared from scratch** are the following.

1) Game Theory Applied to Communications – 2013

- Level: M1 ENSEA
- Type: CM (16 hours) and TP (24 hours)
- Tools: lecture slides, MatLab and the CVX package
- Keywords: convex optimization, descent algorithms (gradient descent, Newton) applied to speech de-noising, classical water-filling algorithm, non-cooperative games, multi-user iterative water-filling algorithm

Year	**New classes**	**Percentage out of 196 hours**
2011 - 2012	212	108.2%
2012 - 2013	59	30.1%
2013 - 2014	45	23%
2014 - 2015	0	0%
2015 - 2016	-	-
2015 - 2017	-	-
2017 - 2018	60	30.6%
2018 - 2019	20	10.2%

Table 1.3: Teaching dynamics: new classes in nb. hours (*nombre d'heures équivalent TD*). The percentages are computed w.r.t. the statutory teaching load of 196 hours per year.

2) Introduction to Game Theory – 2012

- Level: M2R SIC at the University of Cergy-Pontoise
- Type: CM (3 hours)
- Tools: lecture slides
- Keywords: discrete non-cooperative games (prisoner's dilemma, head-tails, etc.), Nash equilibrium, Pareto optimality, von Neumann indifference principle, continuous non-cooperative games, power allocation problem in a distributed multi-user channel

3) Introduction to Digital Communications – 2017

- Level: M1 ENSEA
- Type: CM (8 hours), TD (6 hours) et TP (8 hours)
- Tools: lecture slides, notes, MatLab
- Keywords: analog vs. digital communications, base-band and band-pass communication, basic communication chain, modulation types, inter-symbol interference and Nyquist conditions, adaptive filter, detection

The three modules that, although **existed, I have changed and prepared entirely** are listed below.

4) Advanced Digital Communications – 2012

- Level: M2 ENSEA
- Type: CM (15 hours) et TD (10 hours)
- Tools: lecture notes
- Keywords: high data rate communications, wireless multi-path channel model, equalization, multi-carrier communications OFDM

5) Communication Systems – 2012

- Level: M2 ENSEA
- Type: CM (18 hours) et TD (10 hours)
- Tools: lecture notes
- Keywords: multi-access schemes, OFDMA, SC-FDMA, multi-antenna systems MIMO, spatial diversity, space-time coding (Alamouti)

6) Introduction to C++ Object Oriented Programming – 2011

- Level: M1 ENSEA
- Type: CM (3 hours)
- Tools: lecture slides
- Keywords: encapsulation (classes, objects), inheritance, polymorphism, pointers and references, function and operator overloading, access specifiers

1.1.2 Module Responsibilities

I have been in charge of organising the following modules at ENSEA. This involved: proposing the exam sheets, grading, updating the contents of the modules (CM, TD, TP), etc..

- Game Theory Applied to Communications
- Introduction to Digital Communication
- Advanced Digital Communications
- Communication Systems
- Algorithms
- C++ Object Oriented Programming

1.1.3 Administrative Responsibilities

- During 2014-2015 and 2018-2019, I have been the co-head of the **Networks and Communications Track** - *Option RT - Réseaux et Télécommunications* of the 3rd year of the engineer cycle at ENSEA jointly with Laura Luzzi (ENSEA);
- Since Jun. 2018, I am a jury member of the **Mathematics Admissions Committee of ENSEA**, conducting oral mathematics examinations (*Concours ATS*);
- Since Dec. 2018, I am an elected member of the Technical Committee of ENSEA;
- Since 2013, I have been the ENSEA referee of several engineer internships (M2, M1, L3 levels), e.g., *projet de fin d'études (PFE)*;

- In Nov. 2017, I gave a presentation of ENSEA at the *Lycée Chrestien de Troyes* (*école préparatoire*).

1.2 Research Activities

1.2.1 Summary and Main Research Interests

Since Sep. 2011, my research activities as Associate Professor (*Maître de Conférences*) have been performed mainly at ETIS Lab (*Équipes Traitement de l'Information et Systémes*), UMR 8051, Cergy-Pontoise, France within the communications research group: ICI – *Information, Communication, Imagerie*. In between Sep. 2015 and Aug. 2017, I have been a full-time Visiting Researcher (*délégation compléte*) at Inria Grenoble Rhône-Alpes, France. During this period, I have been associated with two research groups: first with MESCAL (Sep. 2015-Dec. 2015) and then with POLARIS (Jan. 2016-Aug. 2017), which are joint Inria-CNRS project teams and are, hence, also affiliated to LIG (*Laboratoire d'Informatique de Grenoble*), UMR 5217, Grenoble, France.

My research interests are focused on investigating (decision-wise) distributed communication networks composed of intelligent and autonomous nodes (e.g., IoT networks), which potentially may vary in a non-stationary and unpredictable way. The necessary tools to study the interactions between these nodes and the outcomes or network operating points are pluridisciplinary and range from game theory, convex optimization and online optimization to machine learning. The applications I have been investigating stem from: resource allocation problems (spectrum, power, space, time) efficient in terms of rate, power or energy consumption, physical layer security, service pricing problems, state estimation problems in the smart grid, etc..

During my PhD, which was performed in L2S (*Laboratoire des Signaux et Systèmes*) UMR 8506, Gif-sur-Yvette, France under the joint supervision of Samson Lasaulce (CNRS) and Mérouane Debbah (CentraleSupélec, Huawei Paris), my research has been focused on multiple-antenna (MIMO) wireless networks composed of either a single or multiple transmitters communicating to a common receiver (i.e., MAC - multiple access channels). The transmitters were driven by rate maximization individual objectives, in most of the investigated settings, or by energy efficiency maximization. Such problems required the use of convex optimization and non-cooperative games and the developed algorithms leading to the respective (optimal or Nash equilibrium) solutions were based on iterative best-response (leading to iterative water-filling type of) algorithms. Since my PhD, both the range of applications and the range of the exploited theoretical tools have been diversified.

More precisely, since my PhD, I have been interested in the following issues and applications:

- *Design of energy-efficient ressource allocation algorithms for wireless communications*;

Publications: 4 journals [J18-J16], [J13], 1 book chapter [BC4], 11 conf. [C36], [C33-29], [C25-22], [C18], 1 national conf. [CF5]
Collaborators: Panayotis Mertikopoulos (CNRS), Giacomo Bacci (Univ. of Pisa, MBI srl), Luca Sanguinetti (Univ. of Pisa), Romain Negrel (ESIEE), Inbar Fijalkow (ENSEA), Noura Sellami (Univ. of Sfax)
Students: Raouia Masmoudi (PhD), Alexandre Marcastel (PhD), Vineeth Varma (PhD)
Supported by: Inria, Chair Orange IoT, ANR-JCJC-ORACLESS, PEPS CNRS-INS2I JCJC, NEWCOM#, FUI Systematic EconHome, ENSEA

- *Security issues in wireless networks and the use of energy harvesting for as a means against jamming attacks*;

Publications: 2 journals [J20], [J19] and 2 conf. [C35], [C34]
Collaborator: Arsenia Chorti (Univ. of Essex, ENSEA)
Students: Gada Rezgui (MSc), Miroslav Mitev (PhD)
Supported by: Inria, ENSEA

- *Design of rate-efficient ressource allocation algorithms for wireless communications*[1];

Publications: 3 journals [J18], [J11], [J8] and 5 conf. [C32], [C28], [C21], [C16], [C12]
Collaborators: Panayotis Mertikopoulos (CNRS), Aris Moustakas (Kapodistrian Univ. of Athens), Samson Lasaulce (CNRS), Luca Sanguinetti (Univ. of Pisa), Romain Negrel (ESIEE)
Students: Chao Zhang (PhD)
Supported by: Inria, NEWCOM#, ENSEA, L'Oréal France

- *State estimation with privacy constraints in the smart grid;*

Publications: 1 journal [J14] and 2 conf. [C19], [C15]
Collaborators: Lalitha Sankar (Arizona State University) and Vincent H. Poor (Princeton University)
Supported by: Princeton University and Supélec

- *Rate adaptation protocols for video transmission conciliating QoS and user quality of experience (QoE)*;

Publications: 1 journal [J15], 1 conf. [C26]
Collaborator: Ángeles Vasquez-Castro (Universitat Autonoma de Barcelona)
Student: Smrati Gupta (PhD)
Supported by: Universitat Autonoma de Barcelona

- *Subcarrier scheduling in OFDMA systems taking into account the carrier frequency offsets*;

Publications: 1 journal [J12], 1 conf. [C20], 1 national conf. [CF4]
Collaborator: Inbar Fijalkow (ETIS)
Student: Antonia M. Masucci (Post-doc)
Supported by: ENSEA

[1]The journal and conference papers [J18], [C32] appear in both rate- and energy-efficient resource allocation problems since the related work addresses a generic distributed optimization framework that apply to both types of policies.

- *Service pricing techniques in multi-tier heterogeneous networks*;

  Publications: 1 journal [J10] and 2 conf. [C17], [C13]
  Collaborators: Walid Saad (VirginiaTech) and Merouane Debbah (CentraleSupélec, Huawei Paris)
  Students: Camilla M. G. Gussen (MSc), Luca Rose (PhD)
  Supported by: Supélec and Princeton University

- *Hawks and Doves game in a dynamic framework*;

  Publications: 1 journal [J9], 1 conf. [C14]
  Collaborators: Eitan Altman (Inria), Yezekayel Hayel (Univ. d'Avignon)
  Supported by: L'Oréal France.

Most of these issues require the use mathematical tools going beyond classic convex optimization and non-cooperative games (i.e., beyond the tools exploited during my PhD).

- *Generalized non-cooperative games* and *generalized Nash equilibrium* allows to take into account minimum QoS constraints [J13], which implies that not only the objective (or payoff) function of an agent depends on the choices of the others, but also its set of possible choices.

- The *repeated games* framework allows to model repeated interactions as opposed to one-shot interactions and to enable cooperation among selfish agents [J14], [J15].

- At last, *online optimization* and *exponential learning* (inspired from *machine learning* algorithms and *mirror descent* algorithms) are powerful tools to tackle the arbitrary and unpredictable dynamics of a highly varying network [C31], [C30], [C27], [J16], [J11].

- Moreover, *exponential learning algorithms* allow to converge to the Nash equilibrium solution in static multi-user settings in which the iterative water-filling may fail to converge [J16], [J8].

1.2.2 Advising

My overall student advising activities since Sep. 2011 (including the official advising percentages) can be summarized as follows:

- **4 PhD students**: 2 ongoing (75%, 30%) and 2 defended (40%, 50%)

- **2 M2R students**: 150% defended

- **4 M2R research initiation projects**: 350%

- **2 Post-docs**

Ongoing PhD Students

1) Irched Chafaa
   Title: "Resource allocation problems for mmWave systems"
   Period: Nov. 2017-Oct. 2020
   Advising: 75% with Mérouane Debbah (CentraleSupélec, Huawei Paris, 25%, official director)
   Funding: Algerian state PhD fellowship
   Publication: [C38]

2) Kimon Antonalopoulos
   Title: "Learning algorithms for online variational inequalities"
   Period: Nov. 2017-Oct. 2020
   Advising: 30% with Panayotis Mertikopoulos (CNRS, 50%) and Bruno Gaujal (Inria, 20%, official director)
   Funding: Inria PhD fellowship

Defended PhD Students

1) Alexandre Marcastel
   Title: "Online optimization and learning for IoT networks"
   Period: Oct. 2015-Feb. 2019
   Defense date: Feb., 21st 2019
   Advising: 40% with Panayotis Mertikopoulos (CNRS, 30%) and Inbar Fijalkow (ENSEA, 30%, official director)
   Funding: Orange IoT Chair at the University of Cergy-Pontoise Foundation 2015-2018, ATER ENSEA 2018-2019
   Publications: [J21], [C33], [C31], [C30] and a submitted conf. [C39sub]

2) Raouia Masmoudi
   Title: "Home automation communications systems efficient in terms of energy consumption"
   Period: Oct. 2011-Jul. 2015
   Defense date: Dec. 1st, 2015
   Advising: 50% with Inbar Fijalkow (ENSEA, official director)
   Funding: FUI Systematic EconHome 2011-2013, ATER ENSEA 2013-2015
   Publications: [J17], [C36], [C29], [C25], [CF5]

Collaborations involving PhD Students

Throughout my different collaborations, I have worked closely with the PhD students below without being officially involved in their advising, but which have lead to several publications.

1) M. Mitev (University of Essex, UK)
Collaborators: Arsenia Chorti (University of Essex until 2017, ENSEA since 2017) and Martin J. Reed (University of Essex, UK)
Publications: [J22sub]

2) Chao Zhang (CentraleSupélec)
Collaborators: Samson Lasaulce (CNRS)
Publications: [C28]

3) Smrati Gupta (UAB, Spain)
Collaborator: Àngeles Vasquez-Castro (UAB, Spain)
Publications: [J15], [C26]

4) Luca Rose (CentraleSupélec)
Collaborators: Walid Saad (VirginiaTech, USA) and Mérouane Debbah (CentraleSupélec, Huawei, Paris)
Publications: [J10], [C17]

5) Vineeth Varma (CentraleSupélec)
Collaborators: Samson Lasaulce (CNRS) and Mérouane Debbah (CentraleSupélec, Huawei, Paris)
Publications: [BC1]

M2R Students

1) Gada Rezgui
Title: "Energy harvesting as a means to mitigate jamming attacks; a game theoretic analysis"
Defense date: Sep. 14th, 2017
Master: M2R SIC, University of Cergy-Pontoise
Duration: 6 months, 2017
Advising: 50% with Arsenia Chorti (University of Essex, UK)
Publication: [J20]

2) Alexandre Marcastel
Title: "Energy efficient resource allocation policies in cognitive radio systems"
Defense date: Sep. 24th, 2015
Master: M2R SIC, University of Cergy-Pontoise

Duration: 6 months, 2015
Advising: 100%

3) Camilla M. G. Gussen
Title: "Pricing and bandwidth allocation problems in wireless multi-tier networks"
Defense date: 2011
Master: MSc, Universidade Federal do Rio de Janeiro, Brazil
Duration: 4 months, 2011
Advising: 70% with Mérouane Debbah (CentraleSupélec, Huawei Paris, 30%)
Publication: [C13]

M2R Research Initiation Projects

1) Gada Rezgui
Title: "Secret key generation systems under jamming atacks via game theoretic tools"
Master: M2R SIC, University of Cergy-Pontoise
Duration: 3 months, 2017
Advising: 50% with Arsenia Chorti (University of Essex, ENSEA)

2) Alexandre Marcastel
Title: "Interior point algorithms vs. iterative water-filling for resource allocation in cognitive radio systems"
Master: M2R SIC, University of Cergy-Pontoise
Duration: 3 months, 2015
Advising: 100%

3) Arnaud Ngomozogho
Title: "Energy-efficient wireless networks under QoS constaints"
Master: M2R SIC, University of Cergy-Pontoise
Duration: 3 months, 2015
Advising: 100%

4) Dora Boviz
Title: "Comparison of different energy-efficiency measures for wireless communications"
Master: M2R SIC, University of Cergy-Pontoise
Duration: 3 months, 2012
Advising: 100%

Post-docs

1) Olivier Bilenne
Title: "Adaptive optimization algorithms for massive MIMO systems"

Period: Sep. 2018 - Aug. 2019
Collaboration with Panayotis Mertikopoulos (CNRS)
Funding: Project ANR-ORACLESS

2) Antonia M. Masucci
Title: "Subcarrier allocation policies in SC-FDMA systems"
Period: Sep. 2011 - Aug. 2013
Collaboration with Inbar Fijalkow (ENSEA)
Funding: ATER ENSEA
Publications: [J12], [C20], [CF4]

M1 and Undergraduate Students

1) Wang Yuchen
Title: "*Algorithmes d'optimisation de ressources dans les réseaux sans fil distribués*"
Level: M1, ENSEA
Duration: one month and a half, 2017
Advising: 100%

2) Gada Rezgui
Title: "*Algorithmes de stockage distribué dans les réseaux sans fil*"
Level: L3, University of Cergy-Pontoise
Duration: 3 months, 2015
Advising: 50% with Iryna Andriyanova (University of Cergy-Pontoise)

3) Muhammad Abdul Wahab
Title: "*Convergence des algorithmes distribués d'allocation de ressources dans les réseaux sans fils*"
Level: M1, ENSEA
Duration: one month and a half, 2014
Advising: 100%

4) François Denquin
Title: "*Efficacité énergétique dans les réseaux sans fils*"
Level: M1, ENSEA
Duration: one month, 2014
Advising: 100%

1.2.3 Research Projects

Since 2011, I have participated as **Project Member** to the following research projects:

- **Chair Orange IoT** at the University of Cergy-Pontoise Foundation, Oct. 2015 - Sep. 2018, 200k euro
  Collaborators: Inbar Fijalkow (head of the Chair, ENSEA), Jean Schwoerer (Orange) and Alexandre Marcastel (ETIS)
  PhD funding of Alexandre Marcastel (ETIS), 105k euro

- **COST GAMENET** - European Network of Game Theory, CA16228, Oct. 2017 - Sep. 2021
  Collaborators: Mathias Staudigl (Action chair, Maastricht University), Panayotis Mertikopoulos (CNRS)

- **ANR JCJC ORACLESS** - Online resource allocation for unpredictable large-scale wireless systems, Oct. 2016 - Sep. 2021, 207k euro
  Collaborators: Panayotis Mertikopoulos (PI, CNRS), Samson Lasaulce (CNRS)
  Post-doc funding of Olivier Bilenne (LIG), 66k euro

- **PEPS CNRS-INS2I JCJC** - Resource allocation in dynamic network environments via adaptive learning (REAL.net), Jan. 2016 - Dec. 2016, 12.5k euro
  Collaborators: Panayotis Mertikopoulos (PI, CNRS), Alexandre Marcastel (ETIS)

- **ICT NEWCOM #** - Network of Excellence in Wireless Communications, Nov. 2012 - Oct. 2015, funded by the European Commission FP7-ICT-318306
  Participants: CNIT, CNRS, Eurecom, CTTC, Aalborg Universitet, Bilkent Universitesi, IASA, INON, Technion Tel Aviv, Technische Universitaet Dresden, University of Cambridge, Université Catholique de Louvain, Oulun Yliopisto, Technische Universitaet Wien
  Collaborators: Luca Sanguinetti (University of Pisa), Giacomo Bacci (University of Pisa), Panayotis Mertikopoulos (CNRS)

- **FUI Systematic EconHome** - Eco Conception du Home Network, Avr. 2010 - Oct. 2012
  Participants: CEA Grenoble, Docea Power, France Telecom, Orange, INRIA Grenoble, STMicroelectronics
  PhD funding of Raouia Masmoudi (ETIS) during 2011-2013, 65k euro

Since 2013, I have lead as **Principal Investigator** several local projects funded by ENSEA (an overall amount of 10k). Also, I am the **Principal Investigator** of the recently accepted ANR PRCI ELIOT project.

- **ANR PRCI ELIOT** - Enabling Technologies for IoT, **Apr. 2019 - Apr. 2023, 390k euro**
  Role: **Principal Investigator**

International collaboration: ANR-FAPESP (France-Brazil)
Participants: ENSEA, University of São Paulo (USP), and Pontifical Catholic University of Rio de Janeiro (PUC-Rio)
Collaborators: Vitor Nascimento (PI, USP, Brazil), Rodrigo C. de Lamare (PI, PUC-Rio, Brazil), Lukas Landau (PUC-Rio, Brazil), Arsenia Chorti (ENSEA), Iryna Andriyanova (University of Cergy-Pontoise) and Inbar Fijalkow (ENSEA)

- SRV-ENSEA: "Optimisation en TEmps Réel de ressources dans les réseaux mobiles avec Feedback Limité (OTER-FL)", Jan. 2017 - Dec. 2017, 1.2k euro

- SRV-ENSEA, Invited Professor: "Exploiting energy harvesting and spectrum to counteract active jamming attacks", 2017, 1.2k euro
Invited Professor: Arsenia Chorti (University of Essex, UK)

- SRV-ENSEA, Invited Professor: "Game theoretic analysis of physical layer security secret key generation schemes", 2016, 1.2k euro
Invited Professor: Arsenia Chorti (University of Essex, UK)

- SRV-ENSEA: "Optimisation en TEmps Réel de ressources dans les réseaux mobiles qui varient arbitrairement dans le temps (OTER)", Jan. 2016 - Dec. 2016, 1.2k euro

- BQR-ENSEA: "Efficacité énergétique dans les communications 5G (EE5G)", Jan. 2015 - Dec. 2015, 0.6k euro

- BQR-ENSEA, Keynote Speaker: "Wireless Networks: Communication, Cooperation and Competition", Jan. 2014 - Dec. 2014, 1.4k euro
Invited Professor: Zhu Han (University of Houston, USA)

- BQR-ENSEA: "MIMO-MAC Games: MIMO distributed multiple access channel. Game theory and learning", Jan. 2013 - Dec. 2013, 3k euro

1.2.4 Scientific Responsibilities

Editorial Activities

- Since July 2016, I serve as **Executive Editor** for the *Transactions on Emerging Telecommunications Technologies (ETT)*

  Among the Top 11 Editors for outstanding contributions to ETT during the period 2016-2017

- During Feb. 2017 - Oct. 2017, I have served as **Associate Editor** for the *IET Signal Processing* journal.

Conference Organization

- GdR ISIS meeting: "Game Theory, Optimisation and Learning: Interplay and Applications to Signal Processing", Paris, France, 2017
  Co-organized jointly with Samson Lasaulce (CNRS)
  http://www.gdr-isis.fr/index.php?page=reunion&idreunion=346

- WNC3 workshop: "International Workshop on Wireless Networks: Communication, Cooperation and Competition" (in conjunction with WiOpt), Hammamet, Tunisia, 2014
  Co-organized and chaired jointly with Samson Lasaulce (CNRS)
  http://2014wnc3.ensea.fr

- Invited Session: "Communications and Control in the Smart Grid (CCSG)" at IEEE BlackSeaComm, Batumi, Georgia, 2013
  Co-organized jointly with Lalitha Sankar (Arizona State University, USA) and David Gregoratti (CTTC, Spain)

- Session Chair: WCNC (Paris, France) 2012, WNC3 (Hammamet, Tunisia) 2014, GRETSI (Lyon, France) 2015

National Responsibilities

- Jan. 2016 - Dec. 2018, member of the jury of the **Best PhD Prize** in Signal, Image and Computer Vision awarded jointly by the French Club EEA, the GdR ISIS and GRETSI
  - **Vice-president** of the jury in 2017
  - **President** of the jury in 2018

- Jan. 2013 - Nov. 2017, Scientific coordinator (*Directrice Scientifique Adjointe*) of the GdR ISIS, a research group supported by CNRS and industrial partners

  Involved supervising the organization of 4-5 scientific *GdR ISIS meetings* per year jointly with Mérouane Debbah (CentraleSupélec, Huawei Paris, 2013-2014) and Mari Kobayashi (CentraleSupélec, 2015-2017)

  Theme D - Telecommunications: compression, protection, transmission, Axis 2 - Information and communication: from theory to practice

- Examiner of PhD juries
  - Dec. 2nd, 2016 - Examiner of the PhD jury of Kenza Hamidouche, CentraleSupélec, Gif-sur-Yvette
  - Jun. 19th, 2017 - Examiner of the PhD jury of Dora Boviz, Nokia, Paris-Saclay

- Dec. 19th, 2017 - Examiner of the PhD jury of Faton Maliqi, CentraleSupélec, Gif-sur-Yvette
- Dec. 21st, 2017 - Examiner of the PhD jury of Chao Zhang, CentraleSupélec, Gif-sur-Yvette
- Jan. 23rd, 2018 - Examiner of the PhD jury of Philippe Ezran, CentraleSupélec, Gif-sur-Yvette
- Nov. 15th, 2018 - Examiner of the PhD jury of Antony Pottier, IMT Atlantique (previously Télécom Bretagne), Brest
- Dec. 7th, 2018 - Examiner of the PhD jury of Xavier Leturc, Télécom ParisTech

Local Responsibilities

- May 9th, 2017 - Member of an Associate Professor (*Maître de Conférences*) hiring committee at ENSEA
- Sep. 2013 - Mar. 2016 - Elected Member of the Scientific Board of ENSEA

Technical Program Committees (TPC)

- IEEE SPAWC 2019
- IEEE 5G World Forum 2018, 2019
- IEEE PIMRC 2018, 2019
- European Wireless 2018
- IEEE ICC 2012, 2014-2017
- IEEE GLOBECOM 2013-2016
- IEEE WCNC 2012, 2014, 2016-2018
- IEEE ICNC 2013-2019
- IEEE BlackSeaCom 2014-2018
- GameSec 2014
- IEEE INFOCOM 2013
- IEEE SmartGridCom 2013
- ACM ValueTools 2012
- IEEE VTC-Fall 2011

Reviewing

- International journals: *IEEE Trans. on Signal Processing, IEEE Trans. on Inf. Theory, Trans. on Emerging Telecommun. Technologies (ETT), Eurasip JWCN, IEEE Trans on Commun., IEEE Trans. on Wireless Commun., IEEE Trans on Vehicular Technology, IEEE Commun. Lett., IEEE Journal on Sel. Areas in Commun., IEEE Trans. on Smart Grids, IEEE Wireless Commun. Lett., …*
- International conferences: GLOBECOM, ICC, VTC, WiOpt, PIMRC, WCNC, ICNC, ICIP, Gamecomm, Rawnet, WCMC, ISWCS, ValueTools, …
- National conference: GRETSI
- Research projects: French ANR (2013, 2017), European ERC (2016)

1.2.5 International Collaborations

- **Vincent H. Poor**, Princeton University, USA
  Supported by Princeton University, USA and Supélec, France
  Publications: 1 journal + 2 conf.
- **Arsenia Chorti**, University of Essex, UK (until Sep. 2017)
  Supported by ENSEA, France
  Students: Gada Rezgui (M2R) and Miroslav Mitev (PhD)
  Publications: 2 journals + 2 conf.
- **Luca Sanguinetti** and **Giacomo Bacci**, University of Pisa, Italy
  Supported by NEWCOM#
  Publications: 2 journals + 4 conf.
- **Lalitha Sankar**, Arizona State University, USA
  Supported by Princeton University, USA and Supélec
  Publications: 1 journal + 2 conf.
  Conf. organization: Special session at IEEE BlackSeaCom
- **Walid Saad**, VirginiaTech, USA
  Supported by Princeton University and Supélec
  Student: Luca Rose (PhD)
  Publications: 1 journal + 1 conf.
- **Ángeles Vasquez-Castro**, Universitat Autonoma de Barcelona (UAB), Spain
  Supported by UAB, Spain
  Student: Smrati Gupta (PhD)
  Publications: 1 journal + 1 conf.

- **Noura Sellami**, University of Sfax, Tunisia
  Supported by CNRS, France and DGRST, Tunisia
  Student: Raouia Masmoudi (PhD)
  Publications: 3 conf.

- **Aris Moustakas**, National and Kapodistrian University of Athens, Greece
  Supported by L'Oréal
  Publications: 1 journal + 2 conf.

- **Martin J. Reed**, University of Essex, UK
  Student: Miroslav Mitev (PhD)

- **Vitor Nascimento** (USP), **Rodrigo C. De Lamare** (PUC-Rio), **Lukas Landau** (PUC-Rio), Brazil
  Supported by ANR PRCI ELIOT

- **David Gregoratti**, CTTC, Spain
  Conf. organization: Special session at IEEE BlackSeaCom

1.2.6 National Collaborations

- **Panayotis Mertikopoulos**, LIG, CNRS, Grenoble
  Supported by Inria, ANR ORACLESS, PEPS CNRS REAL.net, ENSEA (SRV, BQR), Chair Orange IoT
  Students: Alexandre Marcastel (PhD), Kimon Antonalopoulos (PhD),
  Olivier Bilenne (Post-doc)
  Publications: 5 journals + 9 conf.

- **Mérouane Debbah**, Huawei, Paris
  Supported by CentraleSupélec, Princeton University, Algerian state
  Students: Irched Chafaa (PhD), Luca Rose (PhD), Camilla M. G. Gussen (MSc)
  Conf. organization: GdR ISIS scientific coordination 2013-2014
  Publications (post-PhD): 1 journal + 4 conf.

- **Samson Lasaulce**, L2S, CNRS, Gif-sur-Yvette
  Supported by Supélec, ANR Oracless, L'Oréal, ENSEA, ESIEE
  Conf. organization: WNC3, GdR ISIS meeting
  Publications (post-PhD): 1 journal + 2 conf.

- **Romain Negrel**, ESIEE Paris, Noisy-Champs
  Supported by Inria, ANR Oracless, ESIEE
  Publications: 1 journal

- **Anne Savard**, IMT Douai, Lille
  Publications: 1 conf.
- **Mari Kobayashi**, Supélec, Gif-sur-Yvette
  Conf. organization: GdR ISIS scientific coordination 2015-2017
- **Eitan Altman**, Inria, Avignon
  Supported by L'Oréal
  Publications: 1 journal + 1 conf.
- **Yezekayel Hayel**, Université d'Avignon
  Supported by L'Oréal
  Publications: 1 journal + 1 conf.

1.2.7 Local Collaborations

- **Inbar Fijalkow**, ENSEA
  Supported by EconHome, Chair Orange IoT, ENSEA, CNRS and DGRST, Tunisia, ANR PRCI ELIOT
  Students: Raouia Masmoudi (PhD), Alexandre Marcastel (PhD),
  Antonia M. Masucci (Post-doc)
  Publications: 3 journals + 8 conf.
- **Arsenia Chorti**, ENSEA (since Sep. 2017)
  Supported by ENSEA, ANR PRCI ELIOT
  Students: Gada Rezgui (M2R) and Miroslav Mitev (PhD)
  Publications: 2 journals + 2 conf.
- **Iryna Andriyanova**, Université de Cergy-Pontoise
  Supported by ETIS, ANR PRCI ELIOT
  Student: Gada Rezgui (L3)

1.2.8 Visiting Appointments

- Inria Grenoble Rhône-Alpes, LIG, **France**
  Duration: two years (*délégation*) 2015-2017, one week in 2014, one week in 2013
  Collaborator: Panayotis Mertikopoulos
  Publications: 5 journals + 8 conf.
- Universitat Autonoma de Barcelona, **Spain**
  Duration: one week in 2013
  Collaborator: Ángeles Vasquez-Castro
  Publications: 1 journal + 1 conf.

- Princeton University, New Jersey, **USA**
  Duration: one month in 2012, five months post-doc in 2011
  Collaborators: Vincent H. Poor, Lalitha Sankar, Walid Saad
  Publications: 2 journals + 3 conf.

1.2.9 Invited Talks

- Paris Symposium on Games, Institut Henri Poincaré (IHP), Paris, France, Jun. 2018
  "An application of online mirror descent to wireless communications"
- Transversal Problems in Complexity, Maison Internationale de la Recherche, Cergy-Pontoise, France, May 2018
  "Online mirror descent: An application to wireless communications"
- Lycée Chrestien de Troyes, France, Nov. 2017
  "Introduction à la théorie des jeux et ses applications en communications"
- GDR ISIS Journée Eco Radio, Télécom ParisTech, Paris, France, May 2015 - and - INRIA Rhône Alpes, Grenoble, Oct. 2015
  "Energy-efficient power allocation in dynamic multi-carrier systems"
- CTTC, Barcelona, Spain, Jul. 2013
  "Repeated games for privacy-aware distributed state estimation"
- Workshop on Algorithmic Game Theory: Learning Algorithms and Dynamics in Distributed Systems (AlgoGT), St. Nizier du Moucherotte, France, Jul. 2013
  "Hierarchical games and dynamics in HetNets pricing problems"
- Signal Processing and Optimization for Wireless Communications: **In Memory of Are Hjørungnes** Workshop, Trondheim, Norway, May 2013
  "On a matrix trace inequality"
- GDR ISIS Journée Smart Grids, Télécom ParisTech, Paris, France, Oct. 2011
  "Pricing mechanisms for cooperative state estimation"

1.3 Awards and Distinctions

Post-PhD

2018-2022 Doctoral Supervision and Research Bonus (PEDR) from the French National Council of Universities (CNU 61)

2017 Among the **top 11 Editors** for outstanding contributions to **Transactions on Emerging Telecommunications Technologies (ETT)** during the period 2016-2017

2015-2017 **Inria Fellowship for Mobility** granting two research-dedicated years from Inria, France

2012 **Associate Research Scholar** fellowship from Princeton University, USA funding a one month visit in 2012

Before and during the PhD

2009 **L'Oréal France-UNESCO-French Academy of Science** national fellowship "For young women doctoral candidates in science" (one of the ten laureates), France

2007-2010 **French Ministry of Research and Education** PhD fellowship awarded based on scientific excellence and academic records

2005 **Socrates/Erasmus** international exchange scholarship École Polytechnique, France and Politehnica University of Bucharest, Romania

1.4 Personal Bibliography

The underlined authors (in the publication list) are the names of the PhD and post-doc students that I have co-advised officially and unofficially.

1.4.1 Bibliometry

- Journals: 21 (14 post-PhD) of which 2 (1 post-PhD) invited
- Conferences: 38 (27 post-PhD) of which 11 (4 post-PhD) invited
- Book chapters: 4 (1 post-PhD)
- National conferences: 5 (2 post-PhD)

Index	Total
Citations	1033
h - index	18
i10 - index	25

Table 1.4: Citation indicators: source *Google Scholar*, Jan. 7th, 2019.

1.4.2 Publication Indicators

- Google Scholar webpage:
  https://scholar.google.fr/citations?user=ODy3eccAAAAJ&hl=fr
- HAL webpage (IdHAL elena-veronica-belmega):
  https://cv.archives-ouvertes.fr/elena-veronica-belmega
- ResearchGate webpage:
  https://www.researchgate.net/profile/E_Veronica_Belmega
- ORCID: 0000-0003-4336-4704
- DBLP webpage:
  https://dblp.uni-trier.de/pers/hd/b/Belmega:Elena_Veronica

1.4.3 Peer-reviewed International Journal Papers (post-PhD): 14 + 1 sub.

[J23prep] **E.V. Belmega**, P. Mertikopoulos, R. Negrel and L. Sanguinetti, "Online convex optimization and no-regret learning: Algorithms, guarantees and applications", *in preparation for resubmission to IEEE Signal Processing Magazine*, https://arxiv.org/abs/1804.04529, 2018.

[J22sub] M. Mitev, A. Chorti, **E.V. Belmega**, and M.J. Reed, "Active Attacks in Wireless Secret Key Generation", *submitted to IEEE Commun. Lett.*, Jan. 2019.

[J21] A. Marcastel, **E.V. Belmega**, P. Mertikopoulos, and I. Fijalkow, "Online power optimization in feedback-limited, dynamic and unpredictable IoT networks", ***accepted paper**, IEEE Trans. on Signal Processing*, Mar. 2019.

[J20] G. Rezgui, **E.V. Belmega**, and A. Chorti, "Mitigating jamming attacks using energy harvesting", *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 297 – 300, Feb. 2019.

[J19] **E.V. Belmega**, and A. Chorti, "Protecting secret key generation systems against jamming: Energy harvesting and channel hopping approaches", *IEEE Trans. on Information Forensics & Security*, vol. 12, no. 11, pp. 2611 – 2626, Nov. 2017.

[J18] P. Mertikopoulos, **E.V. Belmega**, R. Negrel, and L. Sanguinetti, "Distributed stochastic optimization via matrix exponential learning", IEEE Trans. on Signal Processing, vol. 65, no. 9, pp. 2277 - 2290, May 2017.

[J17] R. Masmoudi, **E.V. Belmega**, and I. Fijalkow, "Efficient Spectrum Scheduling and Power Management for Opportunistic Users", *EURASIP Journal on Wireless Communications and Networking (JWCN)*, vol. 2016:97, pp. 1 – 19, Apr. 2016.

[J16] P. Mertikopoulos, and **E.V. Belmega**, "Learning to be green: robust energy efficiency maximization in dynamic MIMO-OFDM systems", *IEEE Journal on Selected Areas in Communication, Special Issue on Energy-Efficient Techniques for 5G Wireless Communication Systems*, vol. 34, no. 4, pp. 743 – 757, Apr. 2016.

[J15] S. Gupta, **E.V. Belmega**, and M. A. Vasquez-Castro, "Game theoretical analysis of rate adaptation protocols conciliating QoS and QoE", *EURASIP Journal on Wireless Communications and Networking (JWCN), Special Issue on System Level Modeling in Future Wireless Communications*, vol. 2016:75, pp.1 – 18, Mar. 2016.

[J14] **E.V. Belmega**, L. Sankar and H. V. Poor, "Enabling Data Exchange in Two-Agent Interactive State Estimation under Privacy Constraints", *IEEE Journal of Selected Topics in Signal Processing, Special Issue on Signal and Information Processing for Privacy*, vol. 9, no. 7, pp. 1285 –1297, Oct. 2015.

[J13] G. Bacci, **E.V. Belmega**, P. Mertikopoulos, and L. Sanguinetti, "Energy-Aware Competitive Power Allocation for Heterogeneous Networks Under QoS Constraint", *IEEE Trans. on Wireless Communications*, vol. 14, no. 9, pp. 4728 – 4742, Sep. 2015.

[J12] A.M. Masucci, **E.V. Belmega**, and I. Fijalkow, "Optimal Blockwise Subcarrier Allocation Policies in Single Carrier FDMA Uplink Systems", *EURASIP Journal on Advanced Signal Processing (JASP)*, vol. 2014:176, pp. 1 – 17, Nov. 2014.

[J11] P. Mertikopoulos, and **E.V. Belmega**, "Transmit without Regrets: Online Optimization in MIMO-OFDM Cognitive Radio Systems", *IEEE Journal on Selected Areas in Communications, Cognitive Radio Series*, vol. 32, no. 11, pp. 1987–1999, Nov. 2014.

[J10] L. Rose, **E.V. Belmega**, W. Saad, and M. Debbah, "Pricing in Heterogeneous Wireless Networks: Hierarchical Games and Dynamics", *IEEE Trans. on Wireless Communications*, vol. 13, no. 9, pp. 4985 – 5001, Sep. 2014.

[J9] Y. Hayel, **E.V. Belmega**, and E. Altman, "Hawks and Doves in a dynamic framework", *Dynamic Games and Applications, Springer*, **invited paper**, vol.3, no. 1, pp 24 – 37, Aug. 2012.

[J8] P. Mertikopoulos, **E.V. Belmega**, A. Moustakas, and S. Lasaulce, "Distributed learning policies for power allocation in multiple access channels", *IEEE Journal on Selected Areas in Communications*, vol. 30, no.1, pp. 96 – 106, Jan. 2012.

1.4.4 Peer-reviewed International Journal Papers (PhD): 7

[J7] **E.V. Belmega**, M. Jungers, and S. Lasaulce, "A generalization of a trace inequality for positive definite matrices", *The Australian Journal of Mathematical Analysis and Applications (AJMAA)*, vol. 7, no. 2, pp. 1-5, May 2011.

[J6] **E.V. Belmega** and S. Lasaulce, "Energy-efficient precoding for multiple-antenna terminals", *IEEE. Trans. on Signal Processing*, vol. 59, no. 1, pp. 329–340, Jan. 2011.

[J5] **E.V. Belmega**, B. Djeumou, and S. Lasaulce, "Power allocation games in interference relay channels: Existence analysis of Nash equilibria", *EURASIP Journal on Wireless Communications and Networking (JWCN)*, pp. 120, DOI:10.1155/2010/583462, Nov. 2010.

[J4] **E.V. Belmega**, S. Lasaulce, M. Debbah, M. Jungers, and J. Dumont, "Power allocation games in wireless networks of multi-antenna terminals", *Springer Telecommunications Systems Journal*, vol. 47, no. 1, pp. 109–122, DOI: 10.1007/s11235-010-9305-3, **invited paper**, May 2010.

[J3] **E.V. Belmega**, S. Lasaulce, and M. Debbah, "Power allocation games for MIMO multiple access channels with coordination", *IEEE Trans. on Wireless Communications*, vol. 8, no. 6, pp. 3182–3192, Jun. 2009.

[J2] **E.V. Belmega**, S. Lasaulce, and M. Debbah, "A trace inequality for positive definite matrices", *Journal of Inequalities in Pure and Applied Mathematics (JIPAM)*, vol. 10, no. 1, pp. 1-4, Jan. 2009.

[J1] **E.V. Belmega**, B. Djeumou, and S. Lasaulce, "Gaussian broadcast channels with an orthogonal and bidirectionnal cooperation link", *EURASIP J. on Wireless Communications and Networking (JWCN)*, pp.1–16, doi:10.1155/2008/341726, Apr. 2008.

1.4.5 Book Chapters: 4

[BC4] V. Varma, **E.V. Belmega**, S. Lasaulce, and M. Debbah, "Energy Efficient Communications in MIMO Wireless Channels", *Green Communications: Theoretical Fundamentals, Algorithms, and Applications*, CRC Press, Sep. 2012.

[BC3] **E.V. Belmega**, S. Lasaulce, H. Tembiné, and M. Debbah, "Rate-efficient power allocation games", *Game Theory and Learning for Wireless Networks: Fundamentals and Applications, ISBN 9780123846983*, Academic Press, Elsevier, Jul. 2011.

[BC2] **E.V. Belmega**, S. Lasaulce, and M. Debbah, "Shannon rate efficient power allocation games", *Game Theory for Wireless Communications and Networking*, Auerbach Publications, Taylor and Francis Group, CRC Press, Apr. 2010.

[BC1] **E.V. Belmega**, S. Lasaulce, and M. Debbah, "Capacity of cooperative channels: three terminals case study", *Cooperative Wireless Communication, ISBN 142006469X*, Auerbach Publications, Taylor and Francis Group, CRC Press, Oct. 2008.

1.4.6 Peer-reviewed International Conference Papers (post-PhD): 27 + 1 sub.

[C39sub] A. Marcastel, **E.V. Belmega**, P. Mertikopoulos, and I. Fijalkow, "Gradient-free online resource allocation algorithms for dynamic wireless networks", *submitted paper to IEEE SPAWC 2019*, Feb. 2019.

[C38] I. Chafaa, **E.V. Belmega**, and M. Debbah, "Adversarial Multi-armed Bandit for mmWave Beam Alignment with One-Bit Feedback", *ACM ValueTools 2019*, Palma de Mallorca, Spain, Mar. 2019.

[C37] A. Savard, and **E.V. Belmega**, "Optimal Power Allocation in a Relay-aided Cognitive Networks", *ACM ValueTools 2019*, Palma de Mallorca, Spain, Mar. 2019.

[C36] R. Masmoudi, **E.V. Belmega**, and I. Fijalkow, "Impact of Imperfect CSI on Resource Allocation in Cognitive Radio Channels", *International Workshop on Pervasive and Context-Aware Middleware (PerCAM 17) IEEE WiMOB 2017*, Rome, Italy, Oct. 2017.

[C35] **E.V. Belmega**, and A. Chorti, "Energy Harvesting in Secret Key Generation Systems under Jamming Attacks", *IEEE International Conference on Communications (IEEE ICC)*, Paris, France, May 2017.

[C34] A. Chorti, and **E.V. Belmega**, "Secret Key Generation in Rayleigh Block Fading AWGN Channels under Jamming Attacks", *IEEE International Conference on Communications (IEEE ICC)*, Paris, France, May 2017.

[C33] A. Marcastel, **E.V. Belmega**, P. Mertikopoulos, and I. Fijalkow, "Interference Mitigation via Pricing in Time-Varying Cognitive Radio Systems", **invited paper**, *NetGCoop 2016*, Avignon, France, Nov. 2016.

[C32] P. Mertikopoulos, **E.V. Belmega**, and L. Sanguinetti, "Distributed learning for resource allocation under uncertainty", *IEEE GlobalSIP*, Washington DC, USA, 7-9 Dec. 2016.

[C31] A. Marcastel, **E.V. Belmega**, P. Mertikopoulos, and I. Fijalkow, "Online Interference Mitigation via Learning in Dynamic IoT Environments", *IOE worksop in IEEE GLOBECOM 2016*, Washington DC, USA, 4-8 Dec. 2016.

[C30] A. Marcastel, **E.V. Belmega**, P. Mertikopoulos, and I. Fijalkow, "Online power allocation for opportunistic radio access in dynamic OFDM networks", *IEEE VTC-Fall 2016*, Montreal, Canada, 18-21 Sep. 2016.

[C29] R. Masmoudi, **E.V. Belmega**, I. Fijalkow, and N. Sellami, "Joint scheduling and power allocation in cognitive radio systems", *Advances in Software Defined and Context Aware Cognitive Networks (SCAN) Workshop, IEEE International Conference on Communications (IEEE ICC)*, pp. 399-404, London, UK, 8-12 Jun. 2015.

[C28] C. Zhang, S. Lasaulce, and **E.V. Belmega**, "Using more bandwidth can be detrimental to the global performance in distributed wireless channels", *Small Cell and 5G Networks (SmallNets) Workshop, IEEE International Conference on Communications (IEEE ICC)*, pp. 142-147, London, UK, 8-12 Jun. 2015.

[C27] **E.V. Belmega**, and P. Mertikopoulos, "Learning to be Green: Energy-Efficient Power Allocation in Dynamic Multi-Carrier Systems", *IEEE VTC-Spring*, Glasgow, Scotland, May 2015.

[C26] S. Gupta, **E.V. Belmega**, and M.A. Vasquez-Castro, "Game Theoretical Analysis of the Tradeoff Between QoE and QoS Over Satellite Channels", *7th Advanced Satellite Multimedia Systems Conference 13th Signal Processing for Space Communications Workshop (ASMS/SPSC)*, Livorno, Italy, Sep. 2014.

[C25] R. Masmoudi, **E.V. Belmega**, I. Fijalkow, and N. Sellami, "A unifying view on energy-efficiency metrics in cognitive radio channels", *European Signal Processing Conference (EUSIPCO)*, Lisbon, Portugal, Sep. 2014.

[C24] G. Bacci, **E.V. Belmega**, P. Mertikopoulos, and L. Sanguinetti, "Energy-aware competitive link adaptation in small-cell networks", *The 10th International Workshop on Resource Allocation in Wireless Networks (RAWNET), WiOpt 2014*, **invited paper**, Hammamet, Tunisia, May 2014.

[C23] G. Bacci, **E.V. Belmega**, and L. Sanguinetti, "Distributed energy-efficient power and subcarrier allocation for OFDMA-based small cells", *IEEE International Conf. on Communications (ICC 2014), Workshop on Small Cell and 5G Networks*, Sydney, Australia, Jun. 2014.

[C22] G. Bacci, **E.V. Belmega**, and L. Sanguinetti, "Distributed energy-efficient power optimization in cellular relay networks with minimum rate constraints", *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Florence, Italy, May 2014.

[C21] P. Mertikopoulos, and **E.V. Belmega**, "Distributed Spectrum Management in MIMO-OFDM Cognitive Radio: An Exponential Learning Approach", *ACM International Conference on Performance Evaluation Methodologies and Tools VALUETOOLS 2013*, Torino, Italy, 10-12 Dec. 2013.

[C20] A.M. Masucci, I. Fijalkow, and **E.V. Belmega**, "Subcarrier allocation in coded OFDMA uplink systems: Diversity versus CFO", *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, London, United Kingdom, Sep. 2013.

[C19] **E.V. Belmega**, L. Sankar and H. V. Poor, "Repeated games for privacy-aware distributed state estimation in interconnected networks", *IEEE International Conf. on NETwork Games, COntrol and OPtimization (NETGCOOP)*, Avignon, France, **invited paper**, Nov. 2012.

[C18] R. Masmoudi, **E.V. Belmega**, I. Fijalkow, and N. Sellami, "A Closed-Form Solution to the Power Minimization Problem over Two Orthogonal Frequency Bands under QoS and Cognitive Radio Interference Constraints", *IEEE Dynamic Spectrum Access Networks (DySpan)*, Bellevue, Washington, USA, Oct. 2012.

[C17] L. Rose, **E.V. Belmega**, W. Saad, and M. Debbah, "Dynamic Service Selection Games in Heterogeneous Small Cell Networks with Multiple Providers", *IEEE International Symposium on Wireless Communication Systems (ISWCS)*, Paris, France, Aug. 2012.

[C16] P. Mertikopoulos, **E.V. Belmega**, and A. Moustakas, "Matrix Exponential Learning: Distributed Optimization in MIMO systems", *IEEE International Symposium on Information Theory (ISIT)*, Cambridge, MA, USA, Jun. 2012.

[C15] **E.V. Belmega**, L. Sankar, H. V. Poor, and M. Debbah, "Pricing Mechanisms for Cooperative State Estimation", ISCCSP 2012, Roma, Italy, **invited paper**, May 2012.

[C14] Y. Hayel, **E.V. Belmega**, and E. Altman, "Hawks and Doves in a Dynamic Framework", *International Conf. on NETwork Games, COntrol and OPtimization (NETGCOOP)*, IEEE International Conf. on NETwork Games, COntrol and OPtimization (NETGCOOP), Paris, France, Oct. 2011.

[C13] C.M.G. Gussen, **E.V. Belmega**, and M. Debbah, "Pricing and bandwidth allocation problems in wireless multi-tier networks", *IEEE Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, USA, Nov. 2011.

[C12] P. Mertikopoulos, **E.V. Belmega**, A. Moustakas and S. Lasaulce, "Power Allocation Games in Parallel Multiple Access Channels", *ACM International Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS)*, ENS Cachan, France, May 2011.

1.4.7 Peer-reviewed International Conference Papers (PhD): 11

[C11] **E.V. Belmega**, H. Tembine, and S. Lasaulce, "Learning to precode in outage minimization games over MIMO interference channels", *IEEE Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, USA, **invited paper**, Nov. 2010.

[C10] **E.V. Belmega**, S. Lasaulce, and M. Debbah, "A survey on energy-efficient communications", *IEEE Intl. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC 2010)*, Istanbul, Turkey, **invited paper**, Sep. 2010.

[C9] **E.V. Belmega**, S. Lasaulce, M. Debbah, and A. Hjørungnes, "Learning Distributed Power Allocation Policies in MIMO Channels", *European Signal Processing Conference (EUSIPCO)*, Aalborg, Denmark, **invited paper**, Aug. 2010.

[C8] S. Medina Perlaza, **E.V. Belmega**, S. Lasaulce, and M. Debbah, "On the base station selection and base station sharing in self-configuring networks", *ACM International Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS)*, Pisa, Italy, **invited paper**, Oct. 2009.

[C7] **E.V. Belmega**, and S. Lasaulce, "How useful are multiple antennas in energy-efficient power control games? An information theoretic answer", *ACM International Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS)*, Pisa, Italy, Oct. 2009.

[C6] **E.V. Belmega**, S. Lasaulce, M. Debbah, and A. Hjørungnes "A new energy efficiency function for quasi-static MIMO channels", *International Wireless Communications and Mobile Computing Conference (IWCMC)*, Leipzig, Germany, **invited paper**, Jun. 2009.

[C5] **E.V. Belmega**, B. Djeumou, and S. Lasaulce "Resource allocation games in interference relay channels", *IEEE Intl. Conference on Game Theory for Networks (Gamenets)*, Istanbul, Turkey, **invited paper**, May 2009.

[C4] **E.V. Belmega**, B. Djeumou, and S. Lasaulce "What happens when cognitive terminals compete for a relay node?", *IEEE Intl. Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Taipei, Taiwan, 1–4 Apr. 2009.

[C3] **E.V. Belmega**, S. Lasaulce, and M. Debbah "Power Control in Distributed Multiple Access Channels with Coordination", *IEEE/ACM Proc. of the Intl. Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks and Workshops (WIOPT)*, Berlin, Germany, 1–8 Apr. 2008.

[C2] **E.V. Belmega**, S. Lasaulce, and M. Debbah, "Decentralized handovers in cellular networks with cognitive terminals", *in the IEEE Proc. of the 3rd International Symposium on Communications, Control and Signal Processing (ISCCSP)*, St Julians, Malta, **invited paper**, 12–14 Mar. 2008.

[C1] **E.V. Belmega**, B. Djeumou, and S. Lasaulce, "Performance analysis for the AF-based frequency division cooperative broadcast channel", *in the IEEE Proceedings of the Signal Processing Advances in Wireless Communications conference (SPAWC)*, Helsinki, Finland, 1–5 Jun. 2007.

1.4.8 Peer-reviewed National Conference Papers: 5 + 1 sub.

[CF6] A. Savard, and **E.V. Belmega**, "Allocation de puissance pour les réseaux radio cognitifs à relais", *submitted paper GRETSI*, Lille, France, Mar. 2019.

[CF5] R. Masmoudi, **E.V. Belmega**, I. Fijalkow, and N. Sellami, "Allocation de spectre et de puissance dans les systèmes Radio Cognitives", *GRETSI*, Lyon, France, Sep. 2015.

[CF4] A.M. Masucci, I. Fijalkow, and **E.V. Belmega**, "Politiques optimales d'allocation de blocs de sous-porteuses dans les systèmes OFDMA codés sans connaissance du canal", *GRETSI*, Brest, France, Sep. 2013.

[CF3] B. Djeumou, **E.V. Belmega**, and S. Lasaulce, "Régions de taux atteignables pour le canal à interférence à relais", *GRETSI*, Dijon, France, Sep. 2009.

[CF2] **E.V. Belmega**, B. Djeumou, and S. Lasaulce "Jeux d'allocation de puissance pour les canaux à interférence à relais", *GRETSI*, Dijon, France, Sep. 2009.

[CF1] B. Djeumou, **E.V. Belmega**, and S. Lasaulce, "Recombinaison de signaux décodés et transférés pour le canal à relais à division fréquentielle", *GRETSI*, Troyes, France, 1–4 Sep. 2007.

1.4.9 PhD Thesis

[PHD] **E. V. Belmega**, "On resource allocation problems in distributed MIMO wireless networks (Problémes d'allocation des ressources dans les réseaux MIMO sans fil distribués)", Ph.D. Dissertation, Université Paris Sud-11, 14 December 2010, Gif-sur-Yvette, France, Advisors: S. Lasaulce and M. Debbah.

CHAPTER 2

INTRODUCTION TO ENERGY-EFFICIENT COMMUNICATIONS

Energy efficiency has been identified as one of the major desiderata for the 5th generation communications (5G), and will likely remain a crucial aspect for the information and communications technology (ICT) industry beyond 5G. Energy-efficient communications have been motivating an important part of my research activities during the last four years. In particular, the focus of my work has been on developing resource allocation policies and algorithms that optimize the resource usage (in terms of power, spectrum, space, and time) for energy-efficient data transmission in various wireless networks.

2.1 Why Energy-Efficient Communications?

Owing to the prolific spread of Internet-enabled mobile devices and the ever-growing volume of mobile communication calls, the biggest challenge in the wireless industry today is to meet the soaring demand for wireless broadband required to ensure consistent quality of service (QoS) in a network. Rising to this challenge means increasing the network capacity by a thousandfold over the next few years [1], but the resulting power consumption and energy-related pollution are expected to give rise to major societal, economic and environmental issues that would render this growth unsustainable [2], [3]. Therefore, the ICT industry is faced with a formidable mission: cellular network capacity must be increased significantly in order to accommodate higher data rates, but this task must be accomplished under an extremely tight energy budget. Thus, to achieve the seamless integration of a diverse set of mobile users, applications and services, current design requirements for 5G systems target a dramatic decrease in energy-per-bit consumption of the order of $1,000\times$ or more [1], [4].

A first possible and promising way out of this gridlock is the small-cell (SC) network paradigm which builds on the premise of shrinking wireless cell sizes in order to bring user equipment (UE)

and their serving stations closer to one another. From an operational standpoint, SC networks can be integrated seamlessly into existing macro-cellular networks: the latter ensure wide-area coverage and mobility support, while the former carry most of the generated data traffic [5]. On the other hand, the users' aggressive attitude towards interference from other users can lead to a cascade of power increases at the device level, thereby leading to battery depletion and inefficient energy use. Consequently, solutions that focus exclusively on spectral efficiency maximization are not aligned with energy-efficiency requirements [J6], [C10] – which are crucial for the deployment and operation of heterogeneous networks (HetNets).

A second contending technology to achieve the aforementioned design targets is the emerging massive MIMO (multiple-input and multiple-output) paradigm. Coupled with the use of multiple carrier frequencies via orthogonal frequency division multiplexing (OFDM), massive MIMO "goes large" by employing inexpensive service antennas to focus energy into ever smaller regions of space [6], [7], [8]. As a result, very large MIMO arrays can greatly enhance the reliability of wireless connections and increase throughput and energy efficiency (EE) by a factor of $10\times$ to $100\times$ without requiring the deployment of expensive new air interfaces [8], [9]. However, due to the massive complexity and variability of such systems, a crucial challenge that arises is that wireless users must also be capable of adapting to a dynamic spectrum landscape "on the fly", usually with minimal coordination and limited information at the device end.

Going beyond cellular networks, the Internet of things (IoT) is projected to connect billions of wireless "things" (wireless sensors, wearables, biochip transponders, etc.) in a vast network with drastically different requirements and characteristics between components (e.g., in terms of throughput, power, target applications, etc.) [10]. Following Moore's prediction on silicon integration, the wireless environments of IoT are expected to exhibit massive device densities with high interference levels. An orthogonal spectrum allocation is therefore energetically inefficient, as an unrealistic number of bands or subcarriers would be required to accommodate all devices. The usage of new access protocols such as non-orthogonal multiple access (NOMA) [11] may be considered instead, in which interference mitigation becomes critical. For this reason, and also given that the autonomous wireless devices have stringent battery limitations, optimizing the power consumption emerges as one of the key ingredients for achieving a "speed of thought" user experience at the application level [4].

At last, future devices are likely to be enhanced with many new features among which energy harvesting or scavenging [12] from dedicated RF sources or from the readily available ambient RF radiation could prove essential in boosting the energy autonomy (potentially leading to self-powered devices), in mitigating the network interference or even in mitigating malicious attacks. This has motivated the recent interest surge on RF-based wireless information and power transfer [13], [14], [15], [16], with a focus on wireless sensor networks relevant to IoT applications.

2.2 PhD Earlier Contributions

My PhD has mostly focused on distributed multiple access MIMO systems in which the multiple transmitters compete for the common spectrum and spatial resources and are mainly driven by maximizing their transmission rates at the expense of high power consmumption.

Although energy-efficient communications caught our attention during my PhD [J6], [C10], [C7], [C6] we had solely considered the information-theoretic measure introduced in [17], i.e., the ratio between the Shannon capacity and the power consumption for data transmission, and extended it to a MIMO point-to-point link (composed of a single multi-antenna transmitter and a single multi-antenna receiver). Our investigation in static and fast fading Gaussian channels lead to the extension of the well-known negative result of [17]: in order to maximize the information-theoretic energy efficiency, the transmitter should remain silent and not transmit any data.

[J6] **E.V. Belmega** and S. Lasaulce, "Energy-efficient precoding for multiple-antenna terminals", *IEEE. Trans. on Signal Processing*, vol. 59, no. 1, pp. 329–340, Jan. 2011.

[C10] **E.V. Belmega**, S. Lasaulce, and M. Debbah, "A survey on energy-efficient communications", *IEEE Intl. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC 2010)*, Istanbul, Turkey, **invited paper**, Sep. 2010.

[C7] **E.V. Belmega**, and S. Lasaulce, "How useful are multiple antennas in energy-efficient power control games? An information theoretic answer", *International Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS)*, Pisa, Italy, Oct. 2009.

[C6] **E.V. Belmega**, S. Lasaulce, M. Debbah, and A. Hjørungnes "A new energy efficiency function for quasi-static MIMO channels", *International Wireless Communications and Mobile Computing Conference (IWCMC)*, Leipzig, Germany, **invited paper**, Jun. 2009.

2.3 Post-PhD Contributions

Nevertheless, when accounting for more practical considerations such as including the circuit power consumption [18], [19] or replacing the Shannon data rate by other throughput metrics [20], [21], [22], this negative result no longer holds: a non transmit data to be energy-efficient. My research work post-PhD on this topic goes beyond the point-to-point information-theoretic framework and incorporates diverse aspects such as: studying multi-user interference networks [J18], [J17], [J13], comparing different energy-efficiency metrics [C25], taking into account the dynamics and unpredictability of the networks [J21], [J16], exploiting energy harvesting as a means to counter jamming attacks [J20], [J19].

[J21] A. Marcastel, **E. V. Belmega**, P. Mertikopoulos, and I. Fijalkow, "Online power optimization in feedback-limited, dynamic and unpredictable IoT networks", ***accepted paper***, *IEEE Trans. on Signal Processing*, Mar. 2019.

[J20] G. Rezgui, **E.V. Belmega**, and A. Chorti, "Mitigating jamming attacks using energy harvesting", *accepted paper, IEEE Wireless Commun. Lett.*, Sep. 2018.

[J19] **E.V. Belmega**, and A. Chorti, "Protecting secret key generation systems against jamming: Energy harvesting and channel hopping approaches", *IEEE Trans. on Information Forensics & Security*, vol. 12, no. 11, pp. 2611 – 2626, Nov. 2017.

[J18] P. Mertikopoulos, **E.V. Belmega**, R. Negrel, and L. Sanguinetti, "Distributed stochastic optimization via matrix exponential learning", IEEE Trans. on Signal Processing, vol. 65, no. 9, pp. 2277 - 2290, May 2017.

[J17] R. Masmoudi, **E.V. Belmega**, and I. Fijalkow, "Efficient Spectrum Scheduling and Power Management for Opportunistic Users", *EURASIP Journal on Wireless Communications and Networking (JWCN)*, vol. 2016:97, pp. 1 – 19, Apr. 2016.

[J16] P. Mertikopoulos, and **E.V. Belmega**, "Learning to be green: robust energy efficiency maximization in dynamic MIMO-OFDM systems", *IEEE Journal on Selected Areas in Communication, Special Issue on Energy-Efficient Techniques for 5G Wireless Communication Systems*, vol. 34, no. 4, pp. 743 – 757, Apr. 2016.

[J13] G. Bacci, **E.V. Belmega**, P. Mertikopoulos, and L. Sanguinetti, "Energy-Aware Competitive Power Allocation for Heterogeneous Networks Under QoS Constraint", *IEEE Trans. on Wireless Communications*, vol. 14, no. 9, pp. 4728 – 4742, Sep. 2015.

[BC4] V. Varma, **E.V. Belmega**, S. Lasaulce, and M. Debbah, "Energy Efficient Communications in MIMO Wireless Channels", *Green Communications: Theoretical Fundamentals, Algorithms, and Applications*, CRC Press, Sep. 2012.

[C36] R. Masmoudi, **E.V. Belmega**, and I. Fijalkow, "Impact of Imperfect CSI on Resource Allocation in Cognitive Radio Channels", *International Workshop on Pervasive and Context-Aware Middleware (PerCAM 17) IEEE WiMOB 2017*, Rome, Italy, Oct. 2017.

[C35] **E.V. Belmega**, and A. Chorti, "Energy Harvesting in Secret Key Generation Systems under Jamming Attacks", *IEEE International Conference on Communications (IEEE ICC)*, Paris, France, May 2017.

[C34] A. Chorti, and **E.V. Belmega**, "Secret Key Generation in Rayleigh Block Fading AWGN Channels under Jamming Attacks", *IEEE International Conference on Communications (IEEE ICC)*, Paris, France, May 2017.

[C33] A. Marcastel, **E.V. Belmega**, P. Mertikopoulos, and I. Fijalkow, "Interference Mitigation via Pricing in Time-Varying Cognitive Radio Systems", **invited paper**, *NetGCoop 2016*, Avignon, France, Nov. 2016.

[C32] P. Mertikopoulos, **E.V. Belmega**, and L. Sanguinetti, "Distributed learning for resource allocation under uncertainty", *IEEE GlobalSIP*, Washington DC, USA, 7-9 Dec. 2016.

[C31] A. Marcastel, **E.V. Belmega**, P. Mertikopoulos, and I. Fijalkow, "Online Interference Mitigation via Learning in Dynamic IoT Environments", *IOE worksop in IEEE GLOBECOM 2016*, Washington DC, USA, 4-8 Dec. 2016.

[C30] A. Marcastel, **E.V. Belmega**, P. Mertikopoulos, and I. Fijalkow, "Online power allocation for opportunistic radio access in dynamic OFDM networks", *IEEE VTC-Fall 2016*, Montreal, Canada, 18-21 Sep. 2016.

[C29] R. Masmoudi, **E.V. Belmega**, I. Fijalkow, and N. Sellami, "Joint scheduling and power allocation in cognitive radio systems", *Advances in Software Defined and Context Aware Cognitive Networks (SCAN) Workshop, IEEE International Conference on Communications (IEEE ICC)*, pp. 399-404, London, UK, 8-12 Jun. 2015.

[C27] **E.V. Belmega**, and P. Mertikopoulos, "Learning to be Green: Energy-Efficient Power Allocation in Dynamic Multi-Carrier Systems", *IEEE VTC-Spring*, Glasgow, Scotland, May 2015.

[C25] R. Masmoudi, **E.V. Belmega**, I. Fijalkow, and N. Sellami, "A unifying view on energy-efficiency metrics in cognitive radio channels", *European Signal Processing Conference (EUSIPCO)*, Lisbon, Portugal, Sep. 2014.

[C24] G. Bacci, **E.V. Belmega**, P. Mertikopoulos, and L. Sanguinetti, "Energy-aware competitive link adaptation in small-cell networks", *The 10th International Workshop on Resource Allocation in Wireless Networks (RAWNET), WiOpt 2014*, **invited paper**, Hammamet, Tunisia, May 2014.

[C23] G. Bacci, **E.V. Belmega**, and L. Sanguinetti, "Distributed energy-efficient power and subcarrier allocation for OFDMA-based small cells", *IEEE International Conf. on Communications (ICC 2014), Workshop on Small Cell and 5G Networks*, Sydney, Australia, Jun. 2014.

[C22] G. Bacci, **E.V. Belmega**, and L. Sanguinetti, "Distributed energy-efficient power optimization in cellular relay networks with minimum rate constraints", *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Florence, Italy, May 2014.

[C18] R. Masmoudi, **E.V. Belmega**, I. Fijalkow, and N. Sellami, "A Closed-Form Solution to the Power Minimization Problem over Two Orthogonal Frequency Bands under QoS and Cognitive Radio Interference Constraints", *IEEE Dynamic Spectrum Access Networks (DySpan)*, Bellevue, Washington, USA, Oct. 2012.

[CF5] R. Masmoudi, **E.V. Belmega**, I. Fijalkow, and N. Sellami, "Allocation de spectre et de puissance dans les systèmes Radio Cognitives", *GRETSI*, Lyon, France, Sep. 2015.

My research interests post-PhD have not been exclusively focused on the study of energy-efficient communications, but include various themes (see Sec. 1.2.1 for more details). Since these works and the related contributions go beyond the scope of this manuscript, they will not be presented here.

2.4 Structure of the Remaining Manuscript

The rest of this manuscript is organized as follows. The next chapter is dedicated to the study of energy-efficiency in slowly varying or static networks in an effort to device optimal resource allocation algorithms exploiting tools from convex optimization and non-cooperative game theory (and the different solution concepts: Nash, Debreu, and/or Stackelberg equilibria, depending on the specific problem at hand).

Chapter 4 will incorporate an important ingredient to the energy-efficiency analysis: the temporal variability and unpredictability of the network by focusing on a robust worst-case type of settings, in which there are *no assumptions* on the underlying network dynamics (and can potentially be completely arbitrary and non-stationary). In this context, the objective is to develop adaptive and distributed resource allocation algorithms relying on strictly causal and possibly imperfect feedback information. Such algorithms will draw on online optimization and machine learning tools and will no longer aim at the convergence to a fixed state (such as classic optimal or equilibrium solutions) but meet a different performance criterion in terms of regret minimization.

In Chapter 5, we conclude this HDR thesis by discussing several (short-, mid- and long-term) perspectives and interesting open issues.

At last, Appendix A contains a selection of five publications relevant to energy-efficient communications.

2.5 Notations

Throughout this manuscript we use the following notations: small or large caps x or X for scalar parameters and variables; $\mathbf{x}$ for vectors and $\mathbf{X}$ for matrices. We use the game-theoretic notation (x_k, x_{-k}) when we need to focus on the action x_k of a given player k against the actions of the other players and $x_{-k} = (x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_K)$, where K is the total number of players. At last, we will use of the Landau asymptotic notations $\mathcal{O}(\cdot)$ and $o(\cdot)$. Given two functions $f, g \colon \mathbb{R} \to \mathbb{R}$, we say that "$f$ grows no faster than g asymptotically" and write $f(t) = \mathcal{O}(g(t))$ as $t \to \infty$ if $\limsup_{t\to\infty} f(t)/g(t) < \infty$ or, equivalently, there exists some positive and finite constant $c > 0$ such that $f(t) \leq cg(t)$ for sufficiently large t. Also, we say that "f is dominated by g" and write $f(t) = o(g(t))$, if $\limsup_{t\to\infty} f(t)/g(t) = 0$.

CHAPTER 3

CONTRIBUTIONS EXPLOITING CONVEX OPTIMIZATION AND GAMES

In this chapter, we summarize the main contributions to energy-efficient multi-user communications under the common assumption that the network topology, its characteristics and its parameters remain static throughout the duration of the block transmission time; this means that the channel coherence time is sufficiently large to allow for channel estimation to be performed at the receivers' end (via pilot transmission) and to be fedback to the transmitters. Under this assumption, we derive energy-efficient resource allocation algorithms in different wireless communication settings by exploiting classical tools such as: convex optimization suited for centralized settings, in which a single central authority controls all communication parameters (e.g., the power allocation policies over the available spectrum of the transmitters); and non-cooperative game theory in decentralized settings, in which the communication parameters to be controlled are delegated to each transmitting device.

3.1 Centralized Cognitive Radio Networks

In this section, we overview the most relevant contributions of the work performed by the PhD student Raouia Masmoudi (ETIS), whom I have co-advised (at 50%) jointly with Inbar Fijalkow (ENSEA, 30%, official director) and in collaboration with Noura Sellami (University of Sfax, Tunisia), in the context of the EconHome (FUI Systematic) project leading to the following publications: [J17], [C29], [C25].

More specifically, we study a centralized resource allocation problem in a cognitive radio network designed to address the spectrum scarcity and its current underuse by allowing the opportunistic access to it [23]. A hierarchy between users is imposed, in which secondary users (SUs) are allowed by the spectrum manager to communicate either in the vacant bands left by

the licensed users, called primary users (PUs), or in the non-vacant bands under the condition that the induced interference at the primary receivers is kept below some predefined thresholds.

3.1.1 A Unified View on Energy-Efficiency Metrics

By exploiting the framework of multi-criteria convex optimization [24], we begin by providing a unifying view of three among the most popular energy-efficiency metrics in wireless communications [C25]. Multi-criteria optimization techniques have become popular in wireless communications [25], [26] as they capture the tradeoff operating points between opposing performance criteria. Here, exploiting these tools shows that the different existing energy-efficiency metrics can be unified under the same umbrella. This allows us to compare them and to give insights on choosing the most pertinent metric in a specific scenario. It turns out that the optimal allocation policies maximizing the three energy-efficiency metrics are the Pareto-optimal points lying on the optimal tradeoff curve between the rate maximization and power minimization bi-criteria optimization problem.

We focus on a simple setting composed of a single SU whose objective is to find its most energy-efficient power allocation over the available N frequency sub-carriers while complying with the the constraints imposed by the K PUs. The SU has two opposing desiderata when choosing its best power allocation policy: rate maximization and power consumption minimization. This translates into the following bi-criteria optimization problem:

$$\underset{\mathbf{p}\in\mathscr{P}}{\text{maximize}} \qquad f_o(\mathbf{p}) \triangleq (\, -P_T(\mathbf{p});\;\; C(\mathbf{p})\,) \tag{3.1}$$

where $f_o : \mathbb{R}^N \to \mathbb{R}_- \times \mathbb{R}_+$ is the objective function, $P_T(\mathbf{p})$ denotes the overall transmit power; $C(\mathbf{p})$ is the achievable Shannon rate:

$$P_T(\mathbf{p}) = \sum_{n=1}^{N} p_n, \tag{3.2}$$

$$C(\mathbf{p}) = \sum_{n=1}^{N} \log_2(1+s_n p_n), \tag{3.3}$$

where s_n denotes the quality of the direct link or the effective channel gain incorporating the pathloss, the channel noise and the interference coming from the PUs.

The set $\mathscr{P}$ represents the feasible set shaped by the constraints imposed by the PUs - received peak and average maximum interference ($P_m^{(k)}$ and $\overline{P}^{(k)}$):

$$\mathscr{P} = \left\{ \begin{array}{l|l} \mathbf{p}\in\mathbb{R}_+^N & \sum_{n=1}^{N} g_n^{(k)} p_n \le \overline{P}^{(k)}, \quad \forall\, k, \\ 0 \le g_m^{(k)} p_m \le P_m^{peak(k)}, \;\forall\, k, \forall m & \end{array} \right\}, \tag{3.4}$$

where $g_n^{(k)}$ denotes the channel gain of the interfering link between the k-th PU and the SU in subcarrier n.

There is an inherent conflict among the two desiderata: a) minimizing the power consumption implies a minimum rate equal to zero, i.e., the SU is not transmitting; b) maximizing the rate under the constraints in $\mathscr{P}$ implies a maximum overall power consumption P_T (the rate is a non-decreasing function of the transmit powers).
Therefore, there is no power allocation policy which optimizes both objectives simultaneously; a tradeoff between them has to be made.

The optimal tradeoff points are called Pareto-optimal points that lie on the Pareto-boundary of the set of all feasible power-rate pairs $\mathscr{F}$, defined as:

$$\mathscr{F} = \{(\, P_T(\mathbf{p});\, C(\mathbf{p})\,) \in \mathbb{R}_+ \times \mathbb{R}_+ \mid \mathbf{p} \in \mathscr{P}\}. \tag{3.5}$$

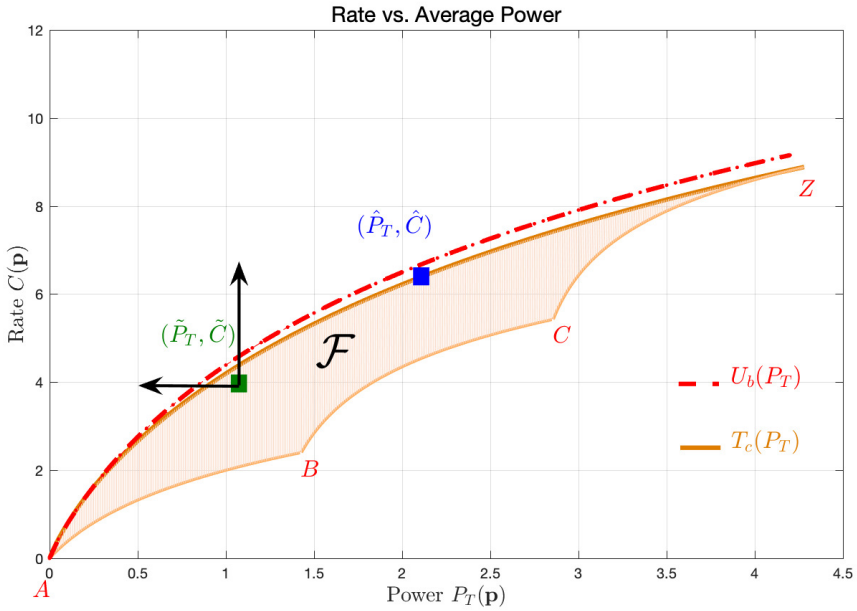


Figure 3.1: The feasible set $\mathscr{F}$ of power-rate pairs in the case of $N = 3$, from a mathematical perspective (from the rate expression in (3.3), we cannot transmit below the Shannon capacity on the worst subcarrier). From an information-theoretic perspective, the entire convex region of power-rate pairs below the Pareto-boundary is achievable (e.g., using standard time-sharing arguments).

We can visualise the set $\mathscr{F}$ in Fig. 3.1, for the following toy setting: $N = 3$ orthogonal frequency bands, $K = 1$ only one PU, power channel gain ordered in decreasing order $\mathbf{s} = [7,\ 5,\ 3]$ and $\mathbf{g} = [7,\ 7,\ 7]$, maximum average interference power constraint $\overline{P} = 40$ and maximum peak interference power constraint $P^{peak(k)} = 10$, $\forall k$. From a purely mathematical perspective, the set

of feasible power-rate pairs is not convex and the peaks B, C are caused by the peak interference constraints. This is because the transmitter is not allowed to transmit below the Shannon capacity, by the definition of $C(\mathbf{p})$. On the contrary, the information-theoretic region is convex and the entire region below the Pareto-boundary is achievable (the transmit rate can be chosen below capacity or by simply considering standard time sharing arguments). Nevertheless, our main interest is the upper-left boundary of $\mathscr{F}$, composed of the tradeoff points $(\widehat{P}_T, \widehat{C})$ called Pareto-optimal and not the interior points $(\widetilde{P}_T, \widetilde{C})$. This boundary is known as the Pareto-boundary and is denoted by $T_c(P_T)$. The upper-bound $U_b(P_T)$ is given by the classical water-filling problem, i.e., rate maximization under overall power constraint P_T:

$$T_c(P_T) = \max\left\{ C(\mathbf{p}) |\ \mathbf{p} \in \mathscr{P},\ \sum_{n=1}^{N} p_n \le P_T \right\} \tag{3.6}$$

$$U_b(P_T) = \max\left\{ C(\mathbf{p}) |\ p_n \ge 0,\ \forall n,\ \sum_{n=1}^{N} p_n \le P_T \right\} \tag{3.7}$$

When operating at a Pareto-optimal point, it is impossible to find any feasible power allocation $\mathbf{p}$ that reduces the power consumption and increases the achievable rate simultaneously. As mentioned above, the pair $(\hat{P}_T, \hat{C})$ in Fig. 3.1 is a Pareto-optimal point, whereas the pair $(\tilde{P}_T, \tilde{C})$ is not. The above problem (3.1) is a convex optimization problem since the objectives are affine and concave and the feasible set is defined by affine inequality constraints. Scalarization, via optimizing the weighted sum of objectives, is a standard technique for finding the Pareto-optimal points in multi-criteria problems [sec.2.6.3 [24]].

Three of the most popular energy-efficiency measures can be unified under the umbrella of the bi-criteria optimization problem in (3.1).

a. *Weighted sum of objectives* [27], [28], [J21]

$$\underset{\mathbf{p} \in \mathscr{P}}{\text{maximize}} \qquad \sum_{n=1}^{N} \log_2(1 + s_n p_n) - \alpha \sum_{n=1}^{N} p_n \tag{3.8}$$

where $\alpha > 0$ can be interpreted as the unitary cost of the power consumption penalizing the rate maximization objective. This approach corresponds to precisely the scalaization of the optimization problem in (3.1).

b. *Power minimization under QoS constraint* [J13], [29], [C18]

$$\begin{aligned} &\underset{\mathbf{p} \in \mathscr{P}}{\text{minimize}} && \sum_{n=1}^{N} p_n \\ &\text{subject to} && C(\mathbf{p}) = \sum_{n=1}^{N} \log_2(1 + s_n p_n) \ge R_{min} \end{aligned} \tag{3.9}$$

where R_{min} denotes the minimum target rate at the SU.

c. *Ratio rate vs. power* [J16], [18], [19]

$$\underset{\mathbf{p} \in \mathscr{P}}{\text{maximize}} \qquad EE(\mathbf{p}) \triangleq \frac{C(\mathbf{p})}{P_T(\mathbf{p}) + P_c} \tag{3.10}$$

where P_c is the fixed circuit power consumption at the SU. This energy-efficiency metric is measured in bits/Joules.

Notice that the solutions of all these three problems are parametrized by α, R_{min} and P_c, respectively and, hence, these parameters have to be very carefully tuned or estimated depending on the application at hand. It turns out that, in all cases, the solutions lie on the Pareto-boundary of the set of feasible power-rate pairs $\mathscr{F}$. However, several differences arise between the three problems.

The first two energy-efficiency problems (3.8) and (3.9) can be solved using classic convex optimization tools and water-filling type of algorithms. Also, all optimal tradeoff pairs (lying on the Pareto-boundary of $\mathscr{F}$) can be easily achieved by tunning either α or R_{min}. At last, there is a geometric relationship between them: if α is the slope of the Pareto-optimal boundary passing through R_{min}, the two problems are equivalent w.r.t. their solutions.

Such a geometric relationship is more difficult to be established between the first two approaches and the ratio rate vs. power maximization in (3.10). Indeed, while the circuit power consumption does not influence the achieved power-rate tradeoff for the first two metrics, this consumption becomes critical for the third one and imposes a minimum level on both the transmission rate and the power consumption for data transmission to operate at optimum. The main difference here is that maximizing the rate vs. power ratio leads to a non-convex optimization problem. Nevertheless, it is a fractional program that can be solved using Dinkelbach iterative method [19], which also explains that the optimal tradeoff point (computed numerically) depending on P_c cannot be identified geometrically on the Pareto-boundary of $\mathscr{F}$.

3.1.2 Efficient Spectrum Scheduling and Power Management for Opportunistic Users

Having discussed three of the most popular energy-efficiency metrics based on the tradeoff between power consumption and Shannon achievable rate, we focus next on the second energy-efficiency metric, i.e., the power consumption minimization under QoS constraint in a multi-user cognitive radio setting [J17]. To be specific, we consider a joint discrete scheduling and power allocation problem at the secondary network from a centralized perspective.

Among the works on resource allocation problems in OFDM systems, without the additional interference constraints of the cognitive radio context, the closest work to ours is [30], in which

a dynamic scheduling and power allocation algorithm was proposed to compute the policies of the multiple non-interfering users that maximize the overall QoS. An algorithm is derived using a Lagrangian relaxation technique to overcome the discrete scheduling constraints. However, a rigorous proof of the convergence and optimality of the proposed algorithm was not provided. Energy-efficiency problems in the cognitive radio context were studied with QoS and spectrum scheduling constraints in [31], in which the authors minimize the SUs' overall power consumption. The framework in [31] is the closest to this work. However, there is no proof of optimality of the proposed scheduling, which is based on a heuristic method involving some exhaustive search steps. In our work, we use convex optimization tools to find the optimal joint scheduling and power allocation under interferences and QoS constraints. Our optimal solution is calculated via an iterative sub-gradient algorithm that is proven to converge to the optimal solution.

The problem under study exhibits two major challenges. First, the minimum QoS and maximum interference constraints may not be simultaneously satisfied and, hence, the problem may not be feasible. To tackle this, we investigate necessary and sufficient conditions for the problem's feasibility. Second, the discrete nature of the channel assignments in the scheduling policy makes the problem non-convex; the centralized authority grants the access to at most one opportunistic user in each subcarrier, to avoid interference imparments. Inspired by [30], we use a Lagrangian relaxation and a dual approach to obtain a solvable convex optimization problem.

Problem formulation

The problem under study can be formalized as follows:

$$
(DP_1)\qquad
\begin{aligned}
&\text{minimize} \sum_{q=1}^{Q}\sum_{n=1}^{N} p_{qn} \\
&\text{s.t.} \quad \sum_{n\in\mathcal{N}} g_{qn}^{(k)} p_{qn} \le \overline{P}_q^{(k)},\ \forall q, \forall k \\
&\qquad 0 \le g_{qn}^{(k)} p_{qn} \le P_{qn}^{peak(k)},\ \forall n, \forall q, \forall k \\
&\qquad C_q(\mathbf{x}_q, \mathbf{p}_q) \ge R_q^{min}, \forall q \\
&\qquad \sum_{q} x_{qn} \le 1, \forall n \\
&\qquad x_{qn} \in \{0,1\}, \forall n, \forall q,
\end{aligned}
$$

where p_{qn} is the power allocated in subcarrier n by SU_q; x_{qn} is the spectrum allocation variable representing whether SU_q is allowed to transmit in subcarrier n; R_q^{min} represents the target QoS at SU_q; $g_{qn}^{(k)}$ is the power gain of the interfering link between SU_q and PU_k; $\overline{P}_q^{(k)}$ is the maximum average interference power that SU_q is allowed to inflict on PU_k; and $P_{qn}^{peak(k)}$ is the maximum peak interference power in subcarrier n that SU_q is allowed to inflict on PU_k. The achievable rate of SU_q is given by

$$
C_q(\mathbf{x}_q, \mathbf{p}_q) = \sum_{n=1}^{N} x_{qn} \log_2\left(1 + s_{qn} p_{qn}\right), \tag{3.11}
$$

where s_{qn} is the effective channel gain of the direct link of SU_q.

As we have already discussed, the above optimization problem is difficult for two reasons. First, the target rate constraints and the maximum interference constraints inflicted on the PUs are opposing ones and, thus, the feasible set may be void depending on the system parameters. Second, to avoid the interference impairments to the PUs and among SUs, we assume that the system owner schedules at most one SU to a given band n and that such a band cannot be further fractioned. This turns the problem into a discrete optimization with respect to the scheduling policy.

Main Contributions

Our main contributions in [J17] (provided in Appendix A) are summarized here below:

- In order to address the feasibility issue, we derive necessary and sufficient conditions on the system parameters for the existence of a solution to the joint spectrum scheduling and power allocation problem in a CR system.
- We formulate a convex optimization problem based on Lagrangian relaxation of the initial non-convex problem. By investigating the Karush-Kuhn-Tucker (KKT) conditions, it turns out that the relaxed problem always has discrete solutions w.r.t. the scheduling policy. These solutions are proved to solve the initial problem (3.11).
- The optimal solution of the relaxed problem, when it exists, is computed via a projected subgradient method [32]. We prove that, when the problem is feasible, our proposed iterative algorithm converges to an optimal solution that satisfies the KKT conditions above (and hence solves the initial problem (3.11)).

Open Issues

Future investigation may include: *i*) studying the robustness of the proposed algorithm to imperfections in the channel state information and, in particular, the issue of underestimating the interfering links to the primary users which may lead to interference constraints violations; *ii*) when the problem is unfeasible, investigating the possibility of selecting a subset of SUs to transmit as opposed to denying the opportunistic access alltogether; *iii*) including a more realistic model in the circuit power consumption, which depends on the transmission data rate, the number of antennas in MIMO systems etc. [33].

3.1.3 Outputs

Regarding the energy-efficiency investigation in centralized cognitive radio networks, the outputs are listed below.

Publications: 1 journal, 4 conf., 1 national conf.

[J17] R. Masmoudi, **E.V. Belmega**, and I. Fijalkow, "Efficient Spectrum Scheduling and Power Management for Opportunistic Users", *EURASIP Journal on Wireless Communications and Networking (JWCN)*, vol. 2016:97, pp. 1 – 19, Apr. 2016.

[C36] R. Masmoudi, **E.V. Belmega**, and I. Fijalkow, "Impact of Imperfect CSI on Resource Allocation in Cognitive Radio Channels", *International Workshop on Pervasive and Context-Aware Middleware (PerCAM 17) IEEE WiMOB 2017*, Rome, Italy, Oct. 2017.

[C29] R. Masmoudi, **E.V. Belmega**, I. Fijalkow, and N. Sellami, "Joint scheduling and power allocation in cognitive radio systems", *Advances in Software Defined and Context Aware Cognitive Networks (SCAN) Workshop, IEEE International Conference on Communications (IEEE ICC)*, pp. 399-404, London, UK, 8-12 Jun. 2015.

[C25] R. Masmoudi, **E.V. Belmega**, I. Fijalkow, and N. Sellami, "A unifying view on energy-efficiency metrics in cognitive radio channels", *European Signal Processing Conference (EUSIPCO)*, Lisbon, Portugal, Sep. 2014.

[C18] R. Masmoudi, **E.V. Belmega**, I. Fijalkow, and N. Sellami, "A Closed-Form Solution to the Power Minimization Problem over Two Orthogonal Frequency Bands under QoS and Cognitive Radio Interference Constraints", *IEEE Dynamic Spectrum Access Networks (DySpan)*, Bellevue, Washington, USA, Oct. 2012.

[CF5] R. Masmoudi, **E.V. Belmega**, I. Fijalkow, and N. Sellami, "Allocation de spectre et de puissance dans les systèmes Radio Cognitives", *GRETSI*, Lyon, France, Sep. 2015.

Collaborators: Inbar Fijalkow (ENSEA), Noura Sellami (Univ. of Sfax, Tunisia)
Student: Raouia Masmoudi (PhD)
Supported by: FUI Systematic EconHome, ENSEA, CNRS, France and DGRST, Tunisia

3.2 Distributed and Autonomous Wireless Networks

In the previous section, a centralized authority was implicitly assumed that is able to compute the power allocation policies of all users and the spectrum scheduling policy. This assumption may be too stringent and unsuitable in various wireless communication settings, such as: a) heterogeneous networks or small cell networks[1], which build on the premise of shrinking wireless cell sizes in order to bring users and their serving stations closer to one another; b) networks suffering from malicious jamming attacks that naturally fall outside of the control of a unique authority; c) underwater acoustic channels, which remain unregulated for the most part [35].

Moreover, distributed resource allocation policies have the important advantage of avoiding the waste of energy associated with centralized algorithms requiring considerable information exchange (and, hence, transmissions) between the users and/or the network administrator [34], [36].

Such settings call for flexible and decentralized resource allocation strategies that rely only on local channel state information (CSI) and require minimal information exchange between network users. This framework is commonly referred to as *distributed optimization*, and it represents a crucial aspect of scalable and efficient network operation. In this section, we investigate solution concepts from *non-cooperative game theory*, an established tool for theoretical tool for problems of this kind.

[1]From an operational standpoint, such networks can be integrated seamlessly into existing macro-cellular networks: the latter ensure wide-area coverage and mobility support, while the former carry most of the generated data traffic [5], [34].

3.2.1 Power Allocation for Heterogeneous Networks under QoS Constraints

We begin by an overview of the work in collaboration with Giacomo Bacci (University of Pisa, Italy), Luca Sanguinetti (University of Pisa, Italy) and Panayotis Mertikopoulos (CNRS) in the context of NEWCOM # - the European Network of Excellence in Wireless Communications, which has lead to the publications [J13], [C23]

Our main goal in this work is the analysis and design of energy-efficient power allocation policies in a heterogeneous network where small cells coexist with macro-tier cellular systems based on orthogonal frequency-division multiple access (OFDMA) technology. Albeit promising, the deployment of small cell networks poses several technical challenges mainly because different small cells are likely to be connected over unreliable infrastructures with widely varying features – such as error rate, outage, delay, and/or capacity specifications. The inherently heterogeneous nature of these networks calls for flexible and decentralized resource allocation strategies that rely only on local CSI and require minimal information exchange between network users and/or access points/base stations.

In particular, focusing on the uplink case, we propose a game-theoretic framework where each user adjusts the allocation of its transmit power (over the available subcarriers) so as to unilaterally maximize its individual link utility subject to a minimum rate requirement. Specifically, each user's energy-aware utility function is defined as the achieved throughput per unit power, accounting for both the power required for data transmission and that required by the circuit components of each user (such as amplifiers, mixer, oscillator, and filters) [19], [33], [C22].

Our work builds on the game-theoretic analysis proposed in [37] where a group of players aims at maximizing their individual energy-efficiency (measured in bits per Watt of transmit power) subject to each user's power constraints. Despite this similarity, the analysis of [37] does not account for minimum rate requirements, thus the resulting game-theoretic model is a standard Nash game with no QoS guarantees – in particular, the users' rates at equilibrium could be fairly low. Incorporating QoS requirements changes the setting drastically and takes us beyond the standard Nash framework because a user's admissible power allocation policy depends crucially on the transmit powers of all other users. The proposed energy-efficient framework represents a generalization of the power minimization under minimum-rate constraints investigated in [29], which is a special case that occurs when the minimum rates are achieved with equality.

Game-theoretic Formulation

Mutual interference in the network introduces interactions among the users that aim at optimizing their energy-efficiency utilities. A natural framework for studying such strategic inter-user interactions is offered by the theory of non-cooperative games with continuous (and action-dependent) action sets. Because of each user's minimum rate constraints, the resulting game departs from the Nash's classical framework [38] and gives rise to a Debreu-type game [39] where

the actions available to each user equipment (UE) depend on the transmit power profile of all other users in the network. In this setting, the relevant solution concept is that of a *DE* – also known as a *generalized Nash equilibrium* [40]. The problem as a non-cooperative game is denoted by $\mathscr{G} \equiv \mathscr{G}(\mathscr{K}, \mathscr{P}, u)$ and consists of the following components:

a) The set of *players* of $\mathscr{G}$ is the set $\mathscr{K}$ of the network's user equipments (UEs).

b) Each player k can choose any transmit power vector in $\mathscr{P}_k^0$ over the N available subcarriers, with

$$\mathscr{P}_k^0 = \left\{ \mathbf{p}_k \in \mathbb{R}_+^N : 0 \leq p_{k,n} \leq \overline{p}_{k,n}, \sum_n p_{k,n} \leq \overline{P}_k \right\} \tag{3.12}$$

for given maximum per-subcarrier transmit power levels $\overline{p}_{k,n}$ and total power constraints $\overline{P}_k$. However, given a power profile $\mathbf{p}_{-k} \in \mathscr{P}_{-k}^0 \equiv \prod_{\ell \neq k} \mathscr{P}_\ell^0$ of the opponents of player k, the *feasible action set* of player k in the presence of the rate requirements is:

$$\mathscr{P}_k(\mathbf{p}_{-k}) = \left\{ \mathbf{p}_k \in \mathscr{P}_k^0 : C_k(\mathbf{p}) \geq \theta_k \right\}. \tag{3.13}$$

where $C_k(\mathbf{p})$ is the achievable rate (normalized to the subcarrier bandwidth, and thus measured in b/s/Hz) of the k-th user:

$$C_k(\mathbf{p}) = \frac{1}{N} \sum_{n=1}^{N} \log_2 \left(1 + \gamma_{k,n}\right) \tag{3.14}$$

with $\gamma_{k,n}$ the signal-to-interference-plus-noise ratio (SINR) over the n-th subcarrier that is achieved by user k at its serving access point (AP):

$$\gamma_{k,n} = \mu_{k,n}(\mathbf{p}_{-k,n})\, p_{k,n} \tag{3.15}$$

$$= \frac{\left| \mathbf{g}_{k,n}^H \mathbf{h}_{kk,n} \right|^2 p_{k,n}}{\|\mathbf{g}_{k,n}\|^2 \sigma^2 + \sum_{j=1, j\neq k}^{K} \left| \mathbf{g}_{k,n}^H \mathbf{h}_{kj,n} \right|^2 p_{j,n}}, \tag{3.16}$$

where $\mathbf{h}_{kj,n}$ $\mathbf{h}_{kj,n} \in \mathbb{C}^{M_{\psi(k)} \times 1}$ denote the uplink channel vector with entries $[\mathbf{h}_{kj,n}]_m$ representing the (frequency) channel gains over subcarrier n from the j-th UE to the m-th receive antenna of the serving AP of user k denoted by $\psi(k)$, where $\psi(k) : \mathscr{K} \mapsto \mathscr{S}$ is a generic function that assigns each user k its serving AP. To keep the complexity at a tolerable level, a simple linear detection scheme is employed for data detection with $\mathbf{g}_{k,n}$ being the vector employed for recovering the data transmitted by user k over subcarrier n.

c) The *utility* $u_k(\mathbf{p}_k; \mathbf{p}_{-k})$ of player k is given by the rate vs. power ratio

$$u_k(\mathbf{p}) = \frac{N^{-1} \sum_{n=1}^{N} \log_2 \left(1 + \mu_{k,n} p_{k,n}\right)}{p_{c,k} + \sum_{n=1}^{N} p_{k,n}}. \tag{3.17}$$

In this framework, the most widely used solution concept is a generalization of the notion of Nash equilibrium [41], known as *Debreu equilibrium* (DE) [39] and sometimes also referred to as *generalized Nash equilibrium* (GNE).

Definition 3.1. A power profile $\mathbf{p}^\star$ is a *Debreu equilibrium* of the energy-efficiency game $\mathcal{G}$ if

$$\mathbf{p}_k^\star \in \mathcal{P}_k(\mathbf{p}_{-k}^\star), \quad \forall k \in \mathcal{K}, \tag{3.18a}$$

and

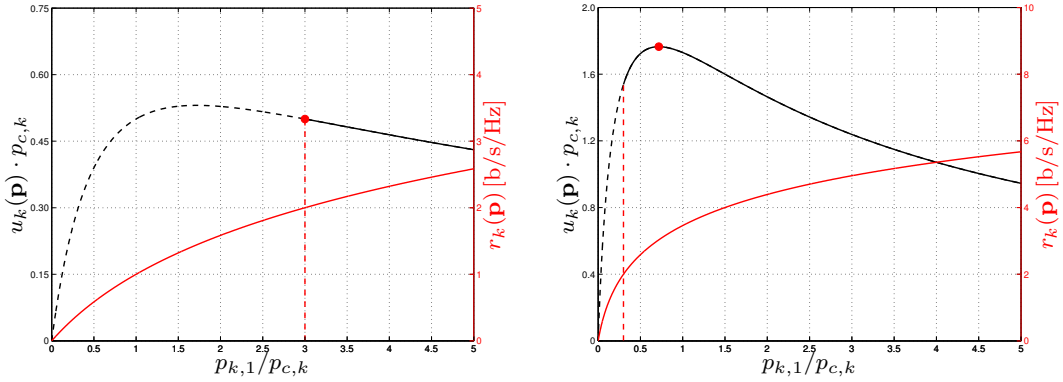
$$u_k(\mathbf{p}^\star) \geq u_k(\mathbf{p}_k; \mathbf{p}_{-k}^\star), \quad \forall \mathbf{p}_k \in \mathcal{P}_k(\mathbf{p}_{-k}^\star),\, k \in \mathcal{K}. \tag{3.18b}$$

The main difference between Debreu and Nash equilibria[2] is that the latter notion considers any unilateral deviations, irrespective of whether the resulting action satisfies the coupled constraints imposed by the actions of other players in the game. Put differently, Nash-type deviations include any action that satisfies a player's individual *uncoupled* constraints, even if so doing violates the player's *coupled* constraints. In the case at hand, this means that, at Nash equilibrium, users would be allowed to transmit at any power level, even if this violates the system's transmission rate requirements. On the other hand, these feasibility constraints are already ingrained in the DE concept: the only unilateral deviations considered are those for which the rate constraints remain satisfied.

As such, Debreu equilibrium are of particular interest in the context of distributed systems because they offer a stable solution of the game from which players (in this case, UEs) have no incentive to deviate and thus destabilize the system, if everyone else maintains their chosen power allocation profiles.

To visualize the impact of the rate constraints on the energy-efficiency optimization, Fig. 3.2 depicts the graph of the utility function (3.17) of user k (normalized by $p_{c,k}$) as a function of the transmit powers $\mathbf{p}_k = \{p_{k,n}\}_{n=1}^N$ for a fixed interference power vector $\mathbf{p}_{-k}$ (and hence keeping $\{\mu_{k,n}(\mathbf{p}_{-k})\}_{n=1}^N$ fixed). We also focus on only one subcarrier, $N = 1$. The dashed black line depicts the unconstrained utility (3.17), whereas the solid black line reports $u_k(\mathbf{p})$ for the values of $p_{k,1}$ such that the rate constraint is met, assuming $\theta_k = 2$ b/s/Hz (for convenience, also the rate r_k is reported with red lines): $\mu_{k,1} = 1/p_{c,k}$ in Fig. 3.2a, whereas $\mu_{k,1} = 10/p_{c,k}$ in Fig. 3.2b. As can be seen, the power level that maximizes $u_k(\mathbf{p})$ (red dot) is on the left boundary of the feasible power set of Fig. 3.2a: in this case, maximizing $u_k(\mathbf{p})$ corresponds to minimizing the power subject to rate constraints, e.g., as considered in [29]. In general however, the maximization of energy efficiency produces a different optimal point, as reported in Fig. 3.2b where the focal user can exploit better channel conditions experienced to increase its utility. This formulation is particularly appealing for next-generation wireless systems, as it captures the tradeoff between obtaining a satisfactory spectral efficiency and saving as much energy as possible [C10], [18], [19].

[2] Recall that a power profile $\mathbf{p}^\star$ is a *Nash equilibrium* of the energy-efficiency game if $u_k(\mathbf{p}^\star) \geq u_k(\mathbf{p}_k; \mathbf{p}_{-k}^\star)$, for all $\mathbf{p}_k \in \mathcal{P}_k^0$ and for all $k \in \mathcal{K}$ (for more details the reader is referred to [41]).



(a) $\mu_{k,1} \cdot p_{c,k} = 1$. (b) $\mu_{k,1} \cdot p_{c,k} = 10$.

Figure 3.2: Normalized utility as a function of the normalized transmit powers ($N = 1$, $\theta_k = 2$ b/s/Hz).

It is easy to see that a particular set of constraints $\{\theta_k\}_{k=1}^K$ may affect the *feasibility* of the problem in the sense that there might not exist *any* power allocation $\mathbf{p} \in \mathbb{R}_+^{K \times N}$ that allows all constraints θ_k to be met *simultaneously* – essentially due to mutual interference in the network, which implies a dependence between the gains $\boldsymbol{\mu}_k \ \forall k$. Necessary and sufficient conditions that ensure the feasibility of the problem without any power constraints and in the single-carrier case $N = 1$ can be found in [C22].

Another problem is that the resulting system could be even *unilaterally infeasible* in the sense that the admissible action set $\mathscr{P}_k(\mathbf{p}_{-k})$ of player k may be empty for a wide range of transmit power profiles $\mathbf{p}_{-k}$ of the other users in the system. Put differently, in the presence of maximum power constraints, any given user may not be able to even participate in the game.

Main Contributions

Our main contributions in [J13] (provided in Appendix A), can be summarized below.

- We characterize the system's Debreu equilibrium as fixed points of a water-filling operator whose water level is a function of the users' minimum rate constraints and circuit power [19] by defining the best-response mapping and using fractional programming techniques [24].

- This characterization is then used to provide sufficient conditions for DE existence and uniqueness and to derive a distributed power allocation algorithm that allows the network to converge to equilibrium under minimal information assumptions. The sufficient conditions ensure that the best-response mapping is a contraction and are somewhat similar in spirit to the well-known conditions ensuring the uniqueness of a Nash equilibrium, in the

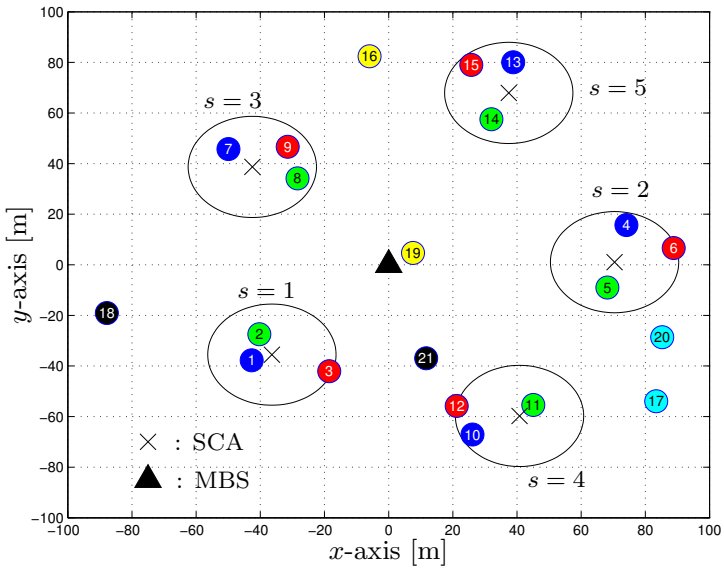


Figure 3.3: Random realization of a network with $S = 5$ small cells, $K_S = 3$ SUEs, and $K_0 = 6$ MUEs, sharing $N = 12$ subcarriers.

non-cooperative rate maximization game studied by [42] in the context of the interference channel. Intuitively, if the interfering connections for a user are sufficiently far away and the resulting SINR is high enough, then the DE exists and is unique.

- The performance of the proposed solution is then validated by means of extensive numerical simulations modeling a HetNet where a macro-tier is augmented with a certain number of low range small-cell access point (SCA). The distributed algorithm we propose is observed to converge to a DE in all the numerical simulations performed and for every network scenario considered, even when the sufficient conditions are not met. As it turns out, the proposed solution represents a scalable and flexible technique to meet the ambitious goals of 5G communications [43], such as extremely high area spectral efficiency (ASE) (more than 500 b/s/Hz/km^2) with a reasonable amount of physical resources (bandwidth and power) and complexity at the network level (number of SCs, signal processing burden, and number of transmit and receive antennas).

To illustrate the performance of the proposed algorithm in a practical setting, Fig. 3.3 reports a random realization of the network with the parameters described in [J13] (see Appendix A). Using the proposed distributed algorithm, after roughly 20 iterations, the users' power profile converges to the DE of $\mathscr{G}$, reported in Fig. 3.4. Here, the first five subplots correspond to the powers allocated in the small cells (the i-th subplot depicts the powers allocated by the users in the si-th small cell with colors matching the ones used in Fig. 3.3), whereas the last two subplots

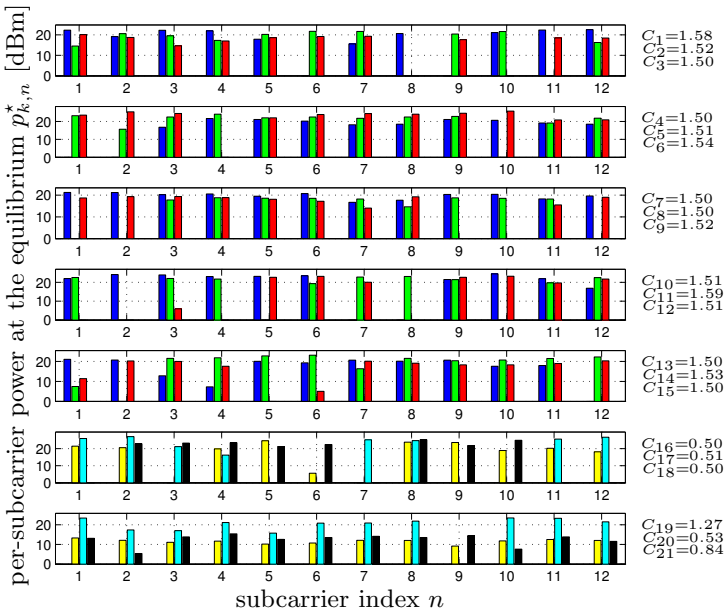


Figure 3.4: Outcome of the proposed resource allocation policy at the DE for the scenario in Fig. 3.3. The subcarriers are allocated exclusively when the MAI within the small cell is large. All users achieve their rate requirements. Users with favorable channels increase their powers to maximize their own utilities.

show the powers selected by the MUEs labeled $\{16,17,18\}$ (in the sixth subplot) and $\{19,20,21\}$ (in the seventh subplot), respectively.

As can be seen in Fig. 3.4, our method tends to allocate the subcarriers in an exclusive manner whenever the multiple access interference (MAI) across UEs within the same small cell is too large (e.g., see the 4th small cell, in which only 5 subcarriers are shared by the 3 users), and to share the same subcarrier when the MAI across users is at a tolerable level (which also includes the interference generated by SUEs from neighboring cells and the MUEs). On the right hand side, we report the achieved rates at the DE in b/s/Hz; all users achieve their minimum demands, while for users with particularly favorable channel conditions (in this case, users no. 1, 11, 19, and 21), it is convenient to increase their transmit power so as to obtain better performance in terms of EE.

As we have already mentioned, we assume the channel to be weakly time-varying and that the convergence of the proposed algorithm is achieved before significant channel variations (common in closed-loop resource allocation schemes). To support this, assume that the uplink and downlink slot durations are in the order of few milliseconds (which is reasonable for LTE/LTE-A standards [44]). In these circumstances, the average convergence time of the proposed solution turns out

to be in the order of tens of milliseconds (since convergence is achieved after approximately 20 iterations): such interval is sufficiently shorter than typical channel coherence times, especially when considering usual SC scenarios with pedestrian users.

Open Issues

Challenging open issues for further work include: *i*) assessing the feasibility of the problem given a particular network realization for the multicarrier case; *ii*) evaluating the impact of different receiver architectures (such as multiuser zero-forcing, and interference cancellation techniques) on the spectral and energy efficiency of the network; *iii*) as an alternative to counter the feasibility issue, investigating the satisfaction equilibrium as a solution concept; *iv*) accounting for highly time-varying scenarios in which users move around the network with high speeds.

3.2.2 Mitigating Jamming Attacks Using Energy Harvesting

In this section, we overview our contributions on the study of a jamming-impaired network composed of a malicious node and a legitimate link with conflicting goals, which naturally falls outside of the centralized optimization framework. This is a joint work with Arsenia Chorti (University of Essex, UK until Sep. 2017; ENSEA from Sep. 2017) and our co-advised (50%-50%) MSc student Gada Rezgui (ETIS), supported locally by ENSEA and ETIS and leading to the publications: [J20], [J19], [C34], [C35].

Secret key generation (SKG) from shared randomness at two remote locations has been extensively studied [45], [46], [47], [48] since the initial works [49], [50], and has been extended to unauthenticated channels [51], [52]. SKG techniques have also been incorporated in protocols that are resilient to spoofing, tampering and man-in-the-middle active attacks [53], [54]. Still, such key generation techniques are not entirely robust against active adversaries, particularly during the advantage distillation phase. Denial of service attacks in the form of jamming are a known vulnerability of SKG systems; in [55], it was demonstrated that when increasing the jamming power, the reconciliation rate normalized to the rate of the SKG increases sharply and the SKG process can in essence be brought to a halt. As SKG techniques are currently being considered for applications such as the IoT [56], the study of appropriate counter-jamming approaches is relevant and timely.

Typically, jamming in wireless communication systems has been investigated using game theoretic tools [57], [58], [59], [60], [61], [62]. Contrary to our work, these earlier studies focus on performance metrics that are either based on the legitimate nodes' signal-to-interference-plus-noise ratio (SINR) [57], [58], [59], [60] and do not incorporate physical-layer security constraints at all, or are based on the secrecy capacity [61], [62]. The secrecy capacity is inherently different than the SKG capacity considered in this work; the former measures the maximum rate at which both confidential and reliable communication is possible, while the latter represents the

maximum rate at which a common secret key can be extracted from the observation of correlated sequences at two remote locations [63].

Next generation terminals are likely to be enhanced with many new features that could prove pivotal in protecting against jamming. For example, greater energy autonomy exploiting energy harvesting (EH) approaches [13], [15] is being researched for systems such as wireless sensor networks for IoT applications. Thus, it is interesting to investigate whether EH could be utilized as a counter-jamming technique by exploiting the harvested jamming power to enhance the quality of the legitimate communication.

To this aim, we study systems in which the legitimate nodes are equipped with EH capabilities and examine whether this added functionality is useful in preempting jamming attacks. We focus on time switching EH protocols [15]: for a fraction of time the legitimate nodes operate in EH mode and switch to the SKG procedure for the rest. The recent work [60] proposes to harvest energy from the jamming interference in a multi-user interference channel in which the jammer is not a strategic decision maker. In terms of formulation, a global optimization problem is investigated (as opposed to an adversarial game). Furthermore, the global performance metric in [60] does not incorporate security constraints and the harvested energy is not directly exploited in the communication phase, appearing only as an additional term in the utility function.

Game-theoretic Formulation

The adversarial interaction between the pair of legitimate nodes and the jammer is formulated as a two-player zero-sum game defined by the tuple $\tilde{\mathcal{G}} = \{\tilde{\mathcal{A}_L}, \tilde{\mathcal{A}_J}, \tilde{u}\left(p,\tau,\gamma\right)\}$ in which the players are: player L representing the legitimate nodes (Alice and Bob act as a single player) on one side, and player J, the jammer, on the other. The action (p,τ) of player L lies in the set $\tilde{\mathcal{A}_L} = [0,P]\times[0,1]$, and the action γ of player J lies in the set $\tilde{\mathcal{A}_J} = [0,\Gamma]$. The objective of player L is to maximize the SKG utility $\tilde{u}(p,\tau,\gamma)$ given as

$$\tilde{u}(p,\tau,\gamma) = \frac{1-\tau}{2}\log_2\left(1+\frac{\left(\frac{p}{1-\tau}+\kappa\gamma\right)\sigma_H^2}{2(1+\sigma^2\gamma)+\frac{(1+\sigma^2\gamma)^2}{\left(\frac{p}{1-\tau}+\kappa\gamma\right)\sigma_H^2}}\right), \tag{3.19}$$

with power constraints $p \le P$, $\gamma \le \Gamma$.

The above expression follows from the fact that we focus on a time-switching EH scheme [15], i.e., we assume that each transmission interval of duration T is divided in two parts. In the first period of duration τT ($0 < \tau \le 1$ being the fraction of T dedicated to EH), both Alice and Bob operate in EH mode with efficiency $0 < \zeta \le 1$; in the second period of duration $(1-\tau)T$, the legitimate nodes operate in SKG mode using the overall available power (including harvested power). For simplicity, we assume that the energy harvested can be stored in a battery without any overflowing issues (unlimited storage) [64]. Furthermore, to ease the mathematical derivation and by ensuring symmetry in the energy harvested at Alice and Bob we assume that $\sigma_A^2 = \sigma_B^2 = \sigma^2$

(the Eve-Alice and Eve-Bob links have equal variance). Given the above considerations and assuming that the energy harvested by Alice and Bob is linear in the received RF power [15], [65]: $E = \zeta\tau T\gamma\sigma^2$, the harvested power for each legitimate node per transmission interval can be expressed as

$$p^{EH} = \frac{E}{(1-\tau)T} = \kappa\gamma, \tag{3.20}$$

where $\kappa = \frac{\zeta\tau\sigma^2}{1-\tau}$ is a convex and increasing function of τ.

The two players are adversaries and the optimal strategy of one player depends on the choice of their opponent such that it cannot be determined unilaterally. In such interactive situations, the Nash equilibrium (NE) [41] is the natural solution concept. Intuitively, a profile $(p^{NE}, \tau^{NE}, \gamma^{NE}) \in \tilde{\mathscr{A}}_L \times \tilde{\mathscr{A}}_J$ is a NE if none of the players can benefit by deviating from this profile knowing that their opponent plays accordingly. Hence, NEs are system states that are stable to unilateral deviations.

We also investigate the so-called Stackelberg equilibrium (SE), which assumes that there is a hierarchy of play (as opposed to the simultaneous play of the NE). The leader of the game, in our case the legitimate player, choses first its optimal action by anticipating the behavior of the follower, i.e., the jammer. The jammer simply observes the choice of the leader and best-responds accordingly [41].

Main Contributions

Our main contributions in [J19] (provided in Appendix A) regarding the game $\tilde{\mathscr{G}}$ can be summarized as follows:

- We reveal the existence of a critical power threshold p_{th} for the legitimate nodes and of an associated threshold harvesting duration τ_{th}. When the legitimate nodes employ EH for longer than τ_{th}, the attacker's optimal strategy is not to jam at all, i.e., the jammer is effectively neutralized. The resulting system state is called the NJ state.

- The Nash and Stackelberg equilibria of the two-player zero-sum game are characterized analytically and in closed form in function of the system parameters.

- We show that neutralizing the jammer is not a stable solution to unilateral deviations (if the strategic decisions are taken simultaneously) and is therefore not a NE of the game. At the NE, it is found that both the legitimate nodes and the jammer transmit with full power and that the energy harvesting (EH) duration does not correspond always to the above threshold. At low signal-to-interference ratio (SIR) (e.g., relatively low transmit power or high jamming power), the EH optimal duration equals τ_{th}. Although the attacker jams with full power, the power collected from EH cancels out the impact of the attack and the SKG capacity is equivalent to the case of using EH for the same duration in absence of

a jammer. At medium to high SIR, the EH optimal duration becomes lower than τ_{th} and decreases until the legitimate nodes do not harvest energy at all.

- Furthermore, when moving to a hierarchical game formulation in which the leader is the the pair of legitimate users and the follower is the jammer, the SE analysis reveals that the legitimate nodes should play the NE strategy. Whenever the legitimate nodes' harvest energy for a duration τ_{th} (at the NE), the jammer neutralization strategy is also a SE solution. This means that, in a hierarchical game, the jammer can potentially be deterred from launching the attack.

In order to compare the SKG capacity at the NE and the NJ state, Fig. 3.5 illustrates the relative gain in utility defined by $E \triangleq \frac{C^{NE}-C^{NJ}}{C^{NE}}$, where $C^{NE} = \tilde{u}(P,\tau^{NE},\Gamma)$ is the utility at the NE and $C^{NJ} = \tilde{u}(p^{NJ},\tau^{NJ},0))$ is the utility at the NJ state, depicted as a function of the SIR: P/Γ for different values of σ^2 and σ_H^2 and for a harvesting efficiency $\zeta = 0.7$. In the investigated settings, the NJ strategy never outperforms the NE in terms of utility, which is consistent with our analysis. When the SIR is relatively low, both the NE and the NJ provide identical utilities. In this case, $p^{NJ} = P$ and $\tau^{NJ} = \tau^{NE} = \tau_{th}(P)$, the jammer is indifferent between $\{0,\Gamma\}$ (as jamming is not harmful when the legitimate user operates at the threshold τ_{th}) and both states are SE solutions. With increasing SIR, it is no longer optimal for the legitimate nodes to harvest energy for a fraction of time $\tau_{th}(P)$ to neutralize the jammer. Instead, by limiting the duration of EH to $\tau^{NE} = \tau_{max} < \tau_{th}(P)$ the SKG capacity increases in spite of the full power jamming $\gamma = \Gamma$ and only the NE is also a SE solution. Moreover, as the SIR increases further, e.g., for $P/\Gamma \gg 1$, the legitimate nodes should not harvest energy at all as the jammer's interference becomes relatively negligible.

Subsequently, we evaluate the impact of the EH capability at the NE. The relative gain in utility obtained at the NE compared with the case in which there is no EH capability ($C^{NoEH} = \tilde{u}(P,0,\Gamma) = C(P,\Gamma)$) is defined as $F \triangleq \frac{C^{NE}-C^{NoEH}}{C^{NE}}$. Fig. 3.6 illustrates F as a function of the SIR: P/Γ, for $\zeta = 0.7$ and various channel parameters. For low SIR, there is a significant gain in utility when employing EH. This gain becomes significantly large at very low SIR, exceeding 97.5 % when the legitimate nodes experience poor channel conditions as opposed to the jammer. When both parties experience similar channel conditions the gain is in the range of 60 % in the medium SIR regime. Overall, the numerical results demonstrate that using EH as a counter-jamming measure is of particular interest in the low and medium SIR regimes but, as expected, it does not increase the utility in the high SIR. The peaks represent here as well the transition from the $\tau^{NE} = \tau_{th}(P)$ regime (at low SIR) to the second regime in which $\tau^{NE} = \tau_{max} < \tau_{th}(P)$.

In our recent work [J20], we further investigate the impact of energy harvesting against jamming attacks when the utility of interest is no longer the SKG capacity but the Shannon achievable rate instead. Similarly, we first demonstrate that the jamming attack can be prevented entirely by adjusting the EH duration, i.e., the jammer can be neutralized (or forced to remain

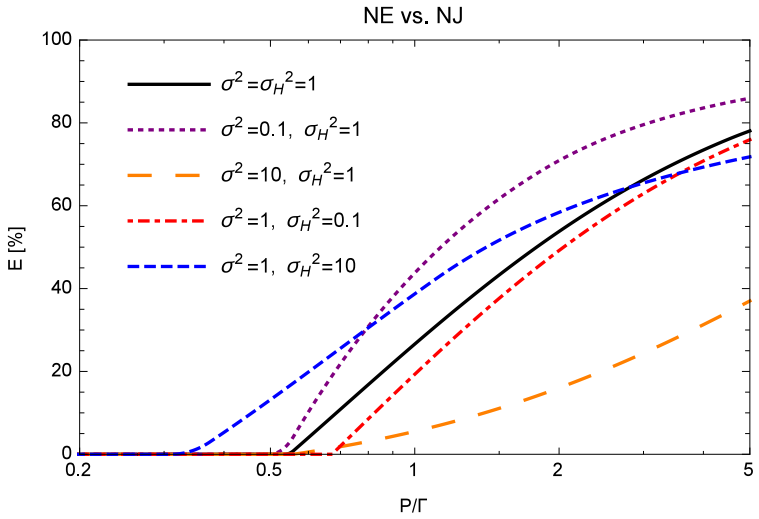


Figure 3.5: Relative utility gain at the NE vs. NJ: $E = \left(C^{NE} - C^{NJ}\right)/C^{NE}$ as a function of $P/\Gamma \geq 0$ for $\zeta = 0.7$.

silent). This is only possible when the quality of the channel in the harvesting link is higher than in the jamming link. Nevertheless, neutralizing the jammer imposes too stringent restrictions on the EH duration and on the legitimate transmit power and hence is not optimal. Second, we formulate a zero-sum game between the legitimate users and the jammer and derive the NE analytically. At the equilibrium, both players transmit at full power, while, the optimal EH duration depends on the system parameters. Interestingly, we show that the NE always outperforms neutralizing the jammer. At the NE, the jamming interference is not fully cancelled but rather exploited, particularly efficient in the high jamming interference regime.

Aside from harvesting the jamming radiated signals, in [J19] we have also considered an alternative way to mitigate jamming attacks by exploiting OFDM systems [58], [66] coupled with channel hopping or power spreading (over the multiple orthogonal subcarriers). Extending the studies in [58], [66] to SKG systems, counter-jamming policies are investigated for N block fading additive white Gaussian noise channels, e.g., systems with N orthogonal subcarriers. At the NE, the jammer always spreads its power over all subcarriers, while for the legitimate nodes the optimality of channel hopping or power spreading depends on the channel parameters. In the high SIR regime, the legitimate nodes should use power spreading to exploit the entire available spectrum given the relatively low jamming interference. On the other hand, at low SIR, the legitimate nodes should use channel hopping and transmit over a single subcarrier to avoid most of the jammer's interference. Furthermore, in characterizing the game's SE we find that the optimal SE strategies reduce to the NE ones, demonstrating that there is no extra payoff to be earned from the advantage of playing first.

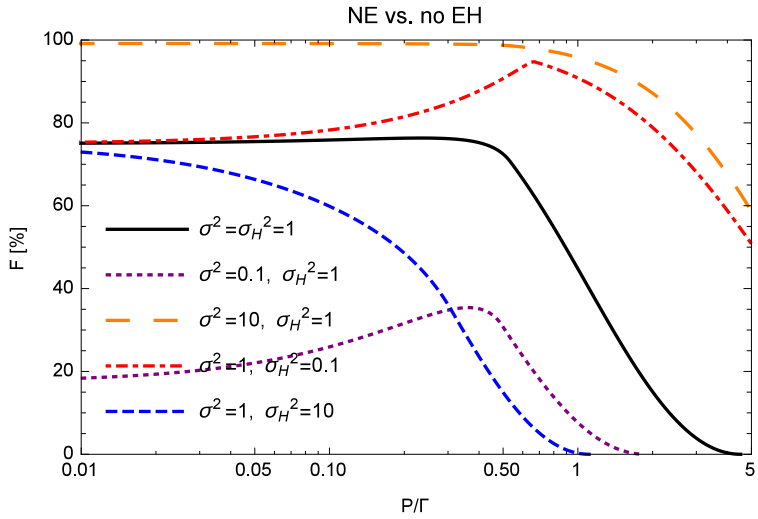


Figure 3.6: Relative utility gain at the NE vs. no EH: $F = \left(C^{NE} - C^{noEH}\right)/C^{NE}$ as a function of $P/\Gamma \geq 0$ for $\zeta = 0.7$ and different channel parameters.

Open Issues

Some interesting extensions for future work are: *i*) investigating the impact of energy harvesting capabilities also at the jammer side - when it can harvest energy from the legitimate transmission and exploit this to enhance the attack; *ii*) studying the secrecy capacity as a utility measure and comparing the results obtained with the three metrics: SKG capacity, Shannon capacity and secrecy capacity; *iii*) replacing the simplified energy harvesting model with a more realistic one that accounts for the wavelength of the RF harvested signals, the distance between an RF energy source and the harvesting node [12]; *iv*) demonstrating the proposed EH policy in a real testbed, in which the effect of imperfect channel estimation, jamming detection errors, implementation aspects of EH, etc., will be considered.

3.2.3 Outputs

The combined outputs of the work on distributed and autonomous wireless networks described in Sec. 3.2.1 (energy-efficient heterogenous networks under QoS constraints) and in Sec. 3.2.2 (mitigating jamming attacks using energy harvesting) are resumed below.

We also mention the joint work with Panayotis Mertikopoulos (CNRS), Romain Negrel (ESIEE) and Luca Sanguinetti (Univ. of Pisa, Italy) published in [J18], [C32], which has not been detailed here. Even though this work falls into the distributed and autonomous networks framework, we address therein a more general problem going beyond energy effciency maximization (the core topic of this thesis).

Publications: 4 journals, 6 conf.

[J20] G. Rezgui, **E.V. Belmega**, and A. Chorti, "Mitigating jamming attacks using energy harvesting", *accepted paper, IEEE Wireless Commun. Lett.*, Sep. 2018.

[J19] **E.V. Belmega**, and A. Chorti, "Protecting secret key generation systems against jamming: Energy harvesting and channel hopping approaches", *IEEE Trans. on Information Forensics & Security*, vol. 12, no. 11, pp. 2611 – 2626, Nov. 2017.

[J18] P. Mertikopoulos, **E.V. Belmega**, R. Negrel, and L. Sanguinetti, "Distributed stochastic optimization via matrix exponential learning", IEEE Trans. on Signal Processing, vol. 65, no. 9, pp. 2277 - 2290, May 2017.

[J13] G. Bacci, **E.V. Belmega**, P. Mertikopoulos, and L. Sanguinetti, "Energy-Aware Competitive Power Allocation for Heterogeneous Networks Under QoS Constraint", *IEEE Trans. on Wireless Communications*, vol. 14, no. 9, pp. 4728 – 4742, Sep. 2015.

[C35] **E.V. Belmega**, and A. Chorti, "Energy Harvesting in Secret Key Generation Systems under Jamming Attacks", *IEEE International Conference on Communications (IEEE ICC)*, Paris, France, May 2017.

[C34] A. Chorti, and **E.V. Belmega**, "Secret Key Generation in Rayleigh Block Fading AWGN Channels under Jamming Attacks", *IEEE International Conference on Communications (IEEE ICC)*, Paris, France, May 2017.

[C32] P. Mertikopoulos, **E.V. Belmega**, and L. Sanguinetti, "Distributed learning for resource allocation under uncertainty", *IEEE GlobalSIP*, Washington DC, USA, 7-9 Dec. 2016.

[C24] G. Bacci, **E.V. Belmega**, P. Mertikopoulos, and L. Sanguinetti, "Energy-aware competitive link adaptation in small-cell networks", *The 10th International Workshop on Resource Allocation in Wireless Networks (RAWNET), WiOpt 2014*, **invited paper**, Hammamet, Tunisia, May 2014.

[C23] G. Bacci, **E.V. Belmega**, and L. Sanguinetti, "Distributed energy-efficient power and subcarrier allocation for OFDMA-based small cells", *IEEE International Conf. on Communications (ICC 2014), Workshop on Small Cell and 5G Networks*, Sydney, Australia, Jun. 2014.

[C22] G. Bacci, **E.V. Belmega**, and L. Sanguinetti, "Distributed energy-efficient power optimization in cellular relay networks with minimum rate constraints", *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Florence, Italy, May 2014.

Collaborators: Giacomo Bacci (Univ. of Pisa, Italy), Luca Sanguinetti (Univ. of Pisa, Italy), Panayotis Mertikopoulos (CNRS), Romain Negrel (ESIEE), Arsenia Chorti (Univ. of Essex, UK, ENSEA)
Student: Gada Rezgui (MSc)
Supported by: Inria, NEWCOM#, PEPS CNRS-INS2I JCJC Real.Net, ENSEA, ETIS, ESIEE

3.3 Conclusions

This chapter provided a concise overview of the main contributions on resource (spectrum, power and time, etc.) allocation policies that optimize the energy efficiency of the communication when the networks are slowly varying in time and when channel state information can be acquired at the transmitting devices' end. The developed algorithms and solutions, either centralized or distributed, rely on techniques from convex optimization and non-cooperative game theory. For each contribution, interesting perspectives and open issues have also been identified.

These methods that target the convergence to a fixed and stable state are not suitable in highly varying networks, which may be unpredictable and possibly non-stationary. Indeed, if by the time the aforementioned algorithms converge to a fixed solution the network state has significantly changed, then these solutions become obsolete. Such arbitrarily varying and unpredictable networks require a new set of tools: online optimization and machine learning, which will be exploited in the next chapter.

CHAPTER 4

CONTRIBUTIONS EXPLOITING ONLINE OPTIMIZATION AND LEARNING

In this chapter, we summarize the main contributions to energy-efficient communications in networks that no longer remain static for the duration of the transmission block. We consider a robust or worse case point of view, in which no assumptions are made regarding the underlying dynamics that governs the temporal variability of the network topology. We start by providing a quick introduction to the online optimization framework, which is a less popular toolbox in the communications community compared to convex optimization and game theory. For a detailed treatment of the topic, the reader is referred to [67] and [68].

4.1 Online Convex Optimization: A Quick Introduction

In classic or static optimization problems, the core underlying assumption is that the objective to be optimized is known by the optimizing agent and remains fixed for the entire runtime of the algorithm computing a solution. Stochastic optimization provides an extension of this framework to problems where the objective function may also depend on a stationary stochastic process. Game theory takes an alternative, multi-agent view of such problems which encourages the optimizing agents to adapt and anticipate changes caused by the other optimizing agents (potentially with conflicting objectives). However, these extensions rely on strong assumptions regarding the variability of the problem's objective, the agents' rationality and common knowledge of rationality (in games), the information at the agents' disposal, etc. By contrast, online optimization provides an elegant toolbox which goes beyond the above by allowing for variations in the problem that are *completely arbitrary* – typically accounting for exogenous (either stationary or not) parameters affecting it, see Fig. 4.1.

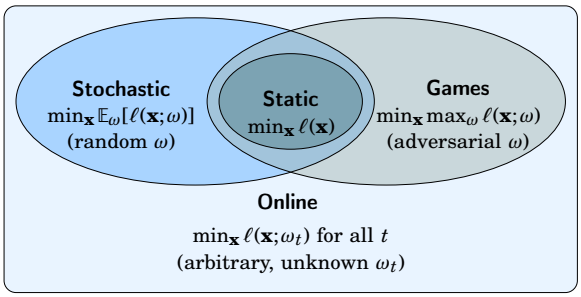


Figure 4.1: A high-level view of the links between online optimization and other frameworks.

The typical setting of online optimization is described as follows. At each stage $t = 1, 2, \ldots$, the optimizing agent selects an action $\mathbf{x}_t$ from some compact convex set $\mathscr{X} \subseteq \mathbb{R}^d$ and incurs a loss of $\ell_t(\mathbf{x}_t)$ based on some loss function $\ell_t : \mathscr{X} \mapsto \mathbb{R}$ that is a priori unknown (this loss function could be determined stochastically, adversarially, or otherwise). Subsequently, the agent selects a new action $\mathbf{x}_{t+1}$ for the next stage and the process repeats as shown below:

Process 1 : Generic online decision process

Require: action set $\mathscr{X}$, sequence of loss functions $\ell_t : \mathscr{X} \to \mathbb{R}$
1: **for** $t = 1$ **to** T **do**
2: play $\mathbf{x}_t \in \mathscr{X}$ # action selection
3: incur $\ell_t(\mathbf{x}_t)$ # incur loss
4: feedback $\mathbf{v}_t$ # receive feedback
5: update $\mathbf{x}_{t+1} \leftarrow \text{function}(\mathbf{x}_t, \mathbf{v}_t)$ # update action based on feedback $\mathbf{v}_t$
6: **end for**

In the above description of online decision processes, no assumptions are made about the structure of the loss functions ℓ_t. The most common settings studied in the related literature [69] are: i) convex online problems, when $\ell_t : \mathscr{X} \to \mathbb{R}$ is convex for all t; and its refinements ii) strongly convex online problems, when $\ell_t : \mathscr{X} \to \mathbb{R}$ is strongly convex for all t; iii) linear online problems, when ℓ_t is of the form $\ell_t(\mathbf{x}) = -\mathbf{v}_t^{\top}\mathbf{x}$ for some vector $\mathbf{v}_t \in \mathbb{R}^d$, t.

Notice that there is a close connection between online optimization and reinforcement learning, a branch of machine learning (along with supervised and unsupervised learning) [69], [70], [71], [72], [73]. In both frameworks, the agent interacts with (or explores) the environment and adapts its action by exploiting observed losses or rewards. On the one hand, the most common online optimization settings mentioned above exploit the overall structure and properties of the problem at hand (w.r.t. the loss functions $\ell_t(\cdot)$ and the feasible action set $\mathscr{X}$) to derive provably efficient decision processes but not on its underlying temporal dynamics and state transitions, which can be completely arbitrary and non-stationary. On the other hand, reinforcement learning (e.g., Q-learning, multi-armed bandit (MAB), Markov decision processes) does not rely on a specific problem model, but may rely on stochastic assumptions on the underlying dynamics and state transitions.

4.1.1 Minimizing the Regret

In the absence of a fixed optimal state to be targeted (given the varying objective in time), the rising questions are: How to evaluate the performance of such an online policy? How to compare two different online policies? A first idea would be to compare the performance in terms of the loss function of an online policy $\mathbf{x}_t$ with the instantaneous optimal solution for all t:

$$\ell_t(\mathbf{x}_t) - \min_{\mathbf{x}\in\mathscr{X}} \ell_t(\mathbf{x}), \quad \forall t \tag{4.1}$$

This comparison is generally too ambitious because of the underlying assumption on the network dynamics, which is varying in a completely arbitrary way and is unknown when the online update $\mathbf{x}_t$ is defined. Instead, a less ambitious and within reach target would be to fix a time horizon T and to compare the overall performance of the online and *dynamic* policy ($\mathbf{x}_t$, $\forall t \leq T$) to the best *fixed* policy in hindsight denoted by $\mathbf{x}^*$:

$$\mathbf{x}^* = \arg\min_{\mathbf{x}\in\mathscr{X}} \sum_{t=1}^{T} \ell_t(\mathbf{x}), \tag{4.2}$$

which minimizes the total incurred loss over the given horizon of play.

Although less ambitious, this fixed policy remains an idealized benchmark that minimizes the total loss incurred over the horizon of play *with perfect knowledge of the future*. This comparison defines the seminal notion of regret introduced in [74], which is the figure of merit to evaluate and compare the performance of online policies.

$$\begin{aligned} \mathrm{Reg}_T &= \sum_{t=1}^{T} \ell_t(\mathbf{x}_t) - \sum_{t=1}^{T} \ell_t(\mathbf{x}^*) && (4.3)\\ &= \sum_{t=1}^{T} \ell_t(\mathbf{x}_t) - \min_{\mathbf{x}\in\mathscr{X}} \sum_{t=1}^{T} \ell_t(\mathbf{x}) && (4.4)\\ &= \max_{\mathbf{x}\in\mathscr{X}} \sum_{t=1}^{T} [\ell_t(\mathbf{x}_t) - \ell_t(\mathbf{x})], && (4.5) \end{aligned}$$

Generally speaking, the regret captures the overall loss incurred by choosing the online *dynamic* policy instead of consistently playing the best *fixed* strategy in hindsight. Since this benchmark requires full, non-causal knowledge of the dynamics governing the evolution of the problem's objective, this is obviously not an implementable policy – it only exists as a theoretical performance target. As such, the aim of online optimization is to derive causal, online learning algorithms that get as close as possible to this remarkable target with the fewest possible assumptions. Consequently, the objective in an online optimization problem is to design a strictly causal, online policy that achieves *no regret*[1], i.e.,

$$\mathrm{Reg}_T = o(T), \tag{4.6}$$

[1] Otherwise stated, the aim is to obtain an average regret that grows sublinearly over the time horizon T: $\limsup_{T\to+\infty} \frac{R_T}{T} \leq 0$ or, equivalently, $\mathrm{Reg}_T = o(T)$.

Algorithm 2 Online gradient descent (OGD)

Require: step-size $\gamma > 0$

choose $\mathbf{x}_1 \in \mathscr{X}$ # initialization

for $t = 1$ **to** T **do**

  incur loss $\hat{\ell}_t \leftarrow \ell_t(\mathbf{x}_t)$ # losses revealed

  observe $\mathbf{v}_t \leftarrow -\nabla \ell_t(\mathbf{x}_t)$ # gradient feedback

  play $\mathbf{x}_{t+1} \leftarrow \Pi(\mathbf{x}_t + \gamma \mathbf{v}_t)$ # gradient step

end for

Online optimization focuses on designing algorithms that attain the best possible regret minimization rate in (4.6), not only in terms of the horizon of play T, but also in terms of the multiplicative constants that come into (4.6), and which depend on the specific problem at hand (i.e., on the geometry and dimensionality of the optimizer's feasible set $\mathscr{X}$ but also on the properties of the objective functions ℓ_t).

4.1.2 Online Learning Algorithms and Their Guarantees

In addition to its wide-ranging scope, another major advantage of the online framework is that the derived algorithms – referred here as *online policies* – come with provable theoretical guarantees in the face of uncertainty. Here we briefly overview two of the most popular online algorithms, which will be exploited later on to design efficient resource allocation algorithms, and their performance in terms of regret.

Online Gradient Descent

The most simple approach for solving classic, offline optimization problems is based on (projected) gradient descent: at each stage, the algorithm takes a step against the gradient of the objective and, if necessary, projects back to the problem's feasible region. Dating back to [75], Online gradient descent (OGD) is the direct adaptation of this idea to an online context and is described in Algorithm 2, where $\Pi\colon \mathbb{R}^d \to \mathscr{X}$ denotes the (Euclidean) projector

$$\Pi(\mathbf{x}') = \operatorname*{argmin}_{\mathbf{x} \in \mathscr{X}} \|\mathbf{x}' - \mathbf{x}\|^2. \tag{4.7}$$

By carefully choosing the step-size γ, the parameter tuning the tradeoff between data exploration and exploitation by controlling the impact of new data and information on the updates, the regret is $\mathrm{Reg}_T = \mathscr{O}(\sqrt{T})$. More precisely, the regret can be bounded as below.

Theorem 4.1. *[75] Against L-Lipschitz convex losses, the OGD algorithm with step-size* $\gamma = (\mathrm{diam}(\mathscr{X})/L)/\sqrt{T}$ *enjoys the regret bound*

$$\mathrm{Reg}_T \leq \mathrm{diam}(\mathscr{X}) L \sqrt{T}, \tag{4.8}$$

with $\mathrm{diam}(\mathscr{X}) \equiv \max_{\mathbf{x},\mathbf{x}' \in \mathscr{X}} \|\mathbf{x}' - \mathbf{x}\|$ *the (Euclidean) diameter of* $\mathscr{X}$.

Algorithm 3 Online mirror descent (OMD)

Require: K-strongly convex regularizer $h : \mathscr{X} \to \mathbb{R}$, step-size $\gamma > 0$

choose $\mathbf{x}_1 \in \mathscr{X}$ # initialization
for $t = 1$ **to** T **do**
  incur loss $\hat{\ell}_t \leftarrow \ell_t(\mathbf{x}_t)$ # losses revealed
  observe $\mathbf{v}_t \leftarrow -\nabla \ell_t(\mathbf{x}_t)$ # gradient feedback
  play $\mathbf{x}_{t+1} \leftarrow P_{\mathbf{x}_t}(\gamma \mathbf{v}_t)$ # mirror step
end for

The performance in terms of regret can be improved under more stringent conditions on the loss functions. In particular, the authors of [76] showed that if the loss functions ℓ_t are strongly convex, the OGD enjoys a logarithmic regret $\mathrm{Reg}_T = \mathscr{O}(\log T)$.

Online Mirror Descent

This second class of algorithms called Online mirror descent (OMD) generalizes OGD and allows to take advantage of the problem's geometry. The main idea is based on the mirror step of [77] for offline problems. Let us re-write the Euclidean projection step of OGD as

$$\begin{aligned}
\mathbf{x}_{t+1} &= \Pi(\mathbf{x}_t + \gamma \mathbf{v}_t) \\
&= \operatorname*{arg\,min}_{\mathbf{x} \in \mathscr{X}} \ \tfrac{1}{2}\|\mathbf{x}_t + \gamma \mathbf{v}_t - \mathbf{x}\|^2 \\
&= \operatorname*{arg\,min}_{\mathbf{x} \in \mathscr{X}} \ \{\tfrac{1}{2}\|\mathbf{x}_t - \mathbf{x}\|^2 + \tfrac{1}{2}\|\gamma \mathbf{v}_t\|^2 + \gamma \mathbf{v}_t^\top (\mathbf{x}_t - \mathbf{x})\} \\
&= \operatorname*{arg\,min}_{\mathbf{x} \in \mathscr{X}} \ \{\gamma \mathbf{v}_t^\top (\mathbf{x}_t - \mathbf{x}) + D(\mathbf{x}_t, \mathbf{x})\}, \qquad (4.9)
\end{aligned}$$

where the term $\frac{1}{2}\|\gamma \mathbf{v}_t\|^2$ is constant w.r.t. $\mathbf{x}$ and where we have defined

$$D(\mathbf{x}', \mathbf{x}) = \tfrac{1}{2}\|\mathbf{x}' - \mathbf{x}\|^2 = \tfrac{1}{2}\|\mathbf{x}'\|^2 - \tfrac{1}{2}\|\mathbf{x}\|^2 - \mathbf{x}^\top (\mathbf{x}' - \mathbf{x}). \qquad (4.10)$$

The key novelty of mirror descent is to replace this quadratic term by the so-called *Bregman divergence*

$$D_h(\mathbf{x}', \mathbf{x}) = h(\mathbf{x}') - h(\mathbf{x}) - \nabla h(\mathbf{x})^\top (\mathbf{x}' - \mathbf{x}), \qquad (4.11)$$

where $h : \mathscr{X} \to \mathbb{R}$ is a smooth K-strongly convex function (usually referred to as a *regularizer*). In so doing, we obtain the OMD algorithm described in Algorithm 3, where the *mirror-prox operator* P is defined as

$$P_{\mathbf{x}'}(\mathbf{v}) = \operatorname*{arg\,min}_{\mathbf{x} \in \mathscr{X}} \ \{\mathbf{v}^\top (\mathbf{x}' - \mathbf{x}) + D_h(\mathbf{x}', \mathbf{x})\}. \qquad (4.12)$$

Before discussing the regret guarantees of the OMD algorithm, it is worth discussing two important special cases.

a) First, as we discussed above, the regularizer $h(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2$ yields the OGD algorithm. As such, all results obtained for OMD also carry over to OGD.

b) The second important instance of OMD is when the problem's feasible region $\mathscr{X}$ is the unit simplex of $\mathbb{R}^d$ and the regularizer is the (negative) Gibbs–Shannon entropy

$$h(\mathbf{x}) = \sum_{j=1}^{d} x_j \log x_j. \tag{4.13}$$

A short calculation shows that the resulting mirror-prox operator is given by the exponential mapping

$$P_{\mathbf{x}'}(\mathbf{v}) = \frac{(x'_j \exp(v_j))_{j=1}^d}{\sum_{k=1}^d x'_k \exp(v_k)} \tag{4.14}$$

which in turn leads to the following algorithm:

$$x_{j,t+1} = \frac{x_{j,t} \exp(\gamma v_{j,t})}{\sum_{k=1}^d x_{k,t} \exp(\gamma v_{k,t})}. \tag{4.15}$$

This algorithm is the well-known *exponential weights* algorithm for MAB with discrete and finite set of actions [78], a fundamental *sequence prediction* (or *forecasting*) problem, which permeates many areas of online and machine learning, and which has found a remarkable breadth of applications – from sparse coding and dictionary learning [79] to filtering [80], matrix prediction [81], channel selection [82], antenna beam selection in mmWave communications [83], network association (access point selection) [84] and many others.

The basic worst-case guarantee of OMD is as follows.

Theorem 4.2. *[67], [68], [85] Against L-Lipschitz convex losses, the OMD algorithm based on a K-strongly convex regularizer h enjoys the regret bound*

$$\mathrm{Reg}_T \le 2L\sqrt{\frac{\max h - \min h}{2K} T}, \tag{4.16}$$

achieved by taking $\gamma = L^{-1}\sqrt{2K(\max h - \min h)/T}$.

OMD vs. OGD

Aside from the fact that OMD generalizes OGD, the former may yield better performance in terms of complexity and decay rate of the regret by exploiting the geometry of the problem at hand.

To be more precise, a first important remark regarding the regret bound above is that the strong convexity and Lipschitz constants K and L need not be taken with respect to the Euclidean norm. In doing so, the regret bound above can be highly improved with respect to the problem's dimensionality (appearing implicitly in the multiplicative constants), which can be a crucial matter in highly dimensional problems (e.g., massive MIMO or Big Data applications).

A second remark is that, depending on the feasible set of the problem under study and on the choice of the regularizer h, the proximal operator may have a closed-form expression as opposed

to the Euclidean projection. This results in a less complex update; as an example consider the unit-simplex and the entropic regularizer for MAB problems above yielding the simple update (4.15).

4.1.3 Outputs

We are currently working on providing a comprehensive overview of the online convex optimization tools and their applications targeting the signal processing and communications communities.

Publication: 1 magazine paper in preparation for resubmission

[J23prep] **E.V. Belmega**, P. Mertikopoulos, R. Negrel and L. Sanguinetti, "Online convex optimization and no-regret learning: Algorithms, guarantees and applications", *in preparation for resubmission to IEEE Signal Processing Magazine*, https://arxiv.org/abs/1804.04529, 2018.

Collaborators: Panayotis Mertikopoulos (CNRS), Romain Negrel (ESIEE), Luca Sanguinetti (Univ. of Pisa, Italy)
Supported by: ANR-JCJC-ORACLESS, Inria, ENSEA

4.2 Self-optimizing and Dynamic Networks

In this section, we overview some applications of the tools described before to energy-efficient resource allocation problems in distributed and autonomous networks (such as ad hoc networks or IoT networks) that vary dynamically over time. Contrary to the static or stationary regime, the networks considered here can evolve in an arbitrary manner, so devices must adapt to such changes on the fly without being able to predict the system state in advance.

4.2.1 Energy Efficiency in Arbitrarily Varying MIMO Networks

We start with the joint work with Panayotis Mertikopoulos (CNRS) [C27], [J16] supported by PEPS CNRS-INS2I JCJC Real.net, ENSEA and Inria, in which we propose a simple and distributed online optimization policy that leads to *no regret*, i.e. it allows users to match (and typically outperform) the best fixed transmit policy in hindsight, irrespective of how the system varies with time. Moreover, to account for the scarcity of perfect CSI in massive MIMO systems, we also study the algorithm's robustness in the presence of measurement errors and observation noise. Importantly, the proposed policy retains its no-regret properties under very mild assumptions on the error statistics and, on average, it enjoys the same performance guarantees as in the noiseless, deterministic case.

Problem Formulation

Consider a distributed wireless network consisting of several point-to-point connections $u \in \mathcal{U} = \{1,\dots,U\}$ (the system's *users*) that are established over a set of orthogonal subcarriers

$s \in \mathscr{S} \equiv \{1,\ldots,S\}$. Each connection $u \in \mathscr{U}$ represents a pair of communicating wireless multi-antenna devices with M_u antennas at the transmitter and N_u antennas at the receiver. Given the distributed nature of the network, we can focus on a specific connection $u \in \mathscr{U}$ and we will treat the multi-user interference vector $\mathbf{w}^s$ as an aggregate noise variable whose covariance depends on the wireless medium and the transmit characteristics of all other users, which will also simplify notations.

Let us denote the *effective channel matrix* of the focal user over subcarrier s by

$$\widetilde{\mathbf{H}}^s = \left(\mathbf{W}^s\right)^{-1/2}\mathbf{H}^s \tag{4.17}$$

The user's Shannon rate (4.18) can be written more concisely as:

$$C(\mathbf{Q}) = \sum\nolimits_{s\in\mathscr{S}} \log\det\left(\mathbf{I} + \widetilde{\mathbf{H}}^s\mathbf{Q}^s(\widetilde{\mathbf{H}}^s)^\dagger\right) = \log\det\left(\mathbf{I} + \widetilde{\mathbf{H}}\mathbf{Q}\widetilde{\mathbf{H}}^\dagger\right), \tag{4.18}$$

where

1. $\mathbf{Q}^s = \mathbb{E}[\mathbf{x}^s(\mathbf{x}^s)^\dagger] \in \mathbb{C}^{M\times M}$ is the user's input signal covariance matrix over subcarrier s.[2]
2. $\mathbf{Q} = \mathrm{diag}(\mathbf{Q}^1,\ldots,\mathbf{Q}^S)$ is the power profile of the focal user over all subcarriers.
3. $\mathbf{W}^s = \mathbb{E}[\mathbf{w}^s(\mathbf{w}^s)^\dagger] \in \mathbb{C}^{N\times N}$ is the multi-user interference covariance matrix of the co-channel interference plus noise affecting the focal connection (obviously, $\mathbf{W}^s$ depends on all other users in the network).

$\widetilde{\mathbf{H}} = \mathrm{diag}(\widetilde{\mathbf{H}}^1,\ldots,\widetilde{\mathbf{H}}^S)$ is the block-diagonal sum of the user's effective channel matrices over all subcarriers. Thus, following [J13], [19], [37], [86], the user's energy efficiency function is defined as his Shannon rate per unit of consumed power, i.e.

$$\mathrm{EE}(\mathbf{Q};t) = \frac{\log\det\left(\mathbf{I} + \widetilde{\mathbf{H}}_t\mathbf{Q}\widetilde{\mathbf{H}}_t^\dagger\right)}{P_c + \mathrm{tr}(\mathbf{Q})} \tag{4.19}$$

where $\mathrm{tr}(\mathbf{Q}) = \sum_s \mathrm{tr}(\mathbf{Q}^s)$ is the user's total transmit power while P_c denotes the total power dissipated in all other circuit components of the transmitting device (mixer, frequency synthesizer, digital-to-analog converter, etc.).

In the expression above, we highlight that $\widetilde{\mathbf{H}}$ collects all sources of noise and interference that cannot be controlled by the focal transmit/receive pair, so the user's energy efficiency objective may vary itself over time in an unpredictable way. On that account, since we wish to focus on dynamic networks that evolve in an arbitrary fashion, we will not be making any specific postulates regarding the behavior of other users in the network and/or the evolution of the user's actual channel matrix $\mathbf{H}$.

This efficiency function (which, formally, has units of bits/Joule) has been widely studied in the literature [J13], [21], [87] and it captures the fundamental trade-off between higher spectral

[2] In the above, expectations are taken over the users' codebooks (assumed Gaussian).

efficiency and increased battery life. An important remark is that the efficiency function depends on the Shannon rate in (4.18). The latter represents the information-theoretic capacity of the focal user's link only under quite stringent assumptions: Gaussian noise, fixed and known effective channel matrix at the receiver side (or Gaussian interference-plus-noise term, and fixed and known channel matrix of the focal link at the receiver), Gaussian and infinite length codebooks, etc.

Most of these information-theoretic assumptions cannot hold in our dynamic setting, in which the connectivity patterns of the users and their channel matrices can vary in an arbitrary (possibly non-stochastic) manner, not to mention the infinite length codebooks - which does not hold even in the least problematic static channel settings. Thus, in our case the Shannon rate is exploited as a utility function and does not represent the capacity in an arbitrarily time-varying channels - a highly nontrivial and open issue. The main reasons behind our choice, justifying also the wide use of this rate approximation, are: a) its simplicity, enabling the devise of tractable and implementable resource allocation policies; b) its relevance to communications, being an increasing function of the useful signal's strength (which conveys the transmitted information) vs. the interference-plus-noise; and c) it allows us to compare our online dynamic policies with existing resource allocation policies that rely on the Shannon rate (e.g., iterative water-filling).

To sum up, in the context of power-limited and energy-aware users, we aim at solving the maximization problem:

$$\begin{aligned} &\text{maximize} \quad \mathrm{EE}(\mathbf{Q};t), \\ &\text{subject to} \quad \mathbf{Q} \in \mathcal{Q}, \end{aligned} \tag{OEE}$$

where

$$\mathcal{Q} = \left\{ \mathrm{diag}(\mathbf{Q}^1, \ldots, \mathbf{Q}^S) : \mathbf{Q}^s \succcurlyeq 0, \textstyle\sum_s \mathrm{tr}(\mathbf{Q}^s) \leq P_{\max} \right\}, \tag{4.20}$$

and $P_{\max}$ denotes the user's maximum transmit power.

Notice the first difficulty is that the online optimization problem (OEE) is an online fractional program and that the user's EE function is not concave. This issue can be overcome by employing the so-called Charnes–Cooper transformation [88] for turning fractional programs into concave ones.

Main Contributions

Our main contributions [J16] can be summarized as follows.

- We exploit the so-called Charnes–Cooper transformation [88] leading to the following variable change

$$\mathbf{X} = \frac{P_c + P_{\max}}{P_{\max}} \frac{\mathbf{Q}}{P_c + \mathrm{tr}(\mathbf{Q})}, \tag{4.21}$$

where we have introduced the normalization constant $(P_c + P_{\max})/P_{\max}$ in order to have $\mathrm{tr}(\mathbf{X}) \leq 1$ for all $\mathbf{Q} \in \mathscr{Q}$ (with equality if and only if $\mathrm{tr}(\mathbf{Q}) = P_{\max}$). Solving for $\mathbf{Q}$ then yields

$$\mathbf{Q} = \frac{P_c P_{\max}}{P_c + P_{\max}(1 - \mathrm{tr}(\mathbf{X}))}\mathbf{X}, \tag{4.22}$$

so, after substitution, we obtain the maximization objective

$$u(\mathbf{X}) = \mathrm{EE}(\mathbf{Q}) = \frac{P_c + P_{\max}(1 - \mathrm{tr}(\mathbf{X}))}{P_c(P_c + P_{\max})} \log\det\left(\mathbf{I} + \frac{P_c P_{\max}\widetilde{\mathbf{H}}\mathbf{X}\widetilde{\mathbf{H}}^{\dagger}}{P_c + P_{\max}(1 - \mathrm{tr}(\mathbf{X}))}\right), \tag{4.23}$$

while the corresponding feasible region attains the simple form:

$$\mathscr{X} = \left\{\mathrm{diag}(\mathbf{X}^1, \ldots, \mathbf{X}^S) : \mathbf{X}^s \succcurlyeq 0 \text{ and } \textstyle\sum_s \mathrm{tr}(\mathbf{X}^s) \leq 1\right\}. \tag{4.24}$$

Given that $C(\mathbf{Q})$ is concave in $\mathbf{Q}$, the function $F(\mathbf{X}, x) = \frac{P_{\max}}{P_c + P_{\max}} x \cdot C(\mathbf{X}/x)$ will be jointly concave in $\mathbf{X}$ and x, so $u(\mathbf{X})$ will also be concave in $\mathbf{X}$ as the restriction of $F(\mathbf{X}, x)$ to the convex set $P_c P_{\max} x = P_c + P_{\max}(1 - \mathrm{tr}(\mathbf{X}))$ [24]. In this way, (OEE) boils down to the online concave maximization problem:

$$\begin{aligned} &\text{maximize} \quad u(\mathbf{X}; t), \\ &\text{subject to} \quad \mathbf{X} \in \mathscr{X}, \end{aligned} \tag{4.25}$$

where, as before, the dependence on $t = 1, 2 \ldots$, reflects the evolution of the user's effective channel matrices over time.

- We then propose a no-regret transmit policy $\mathbf{X}(t)$ based on OGD described in Sec. 4.1.2 and detailed in Algorithm 4[3] for the online concave problem (4.25) and we will then use the inverse transformation (4.22) to obtain a no-regret policy for (OEE). At each iteration, the knowledge of the gradient matrix $\mathbf{V} = \nabla u$ required and assumed to be bounded $\|\mathbf{V}\| \leq V$ with

$$\mathbf{V} = \nabla u = \frac{P_{\max}}{P_c + P_{\max}}\left[\mathbf{A} + \frac{\mathrm{tr}(\mathbf{A}\mathbf{Q}) - C(\mathbf{Q})}{P_c} \cdot \mathbf{I}\right], \tag{4.26}$$

where $\mathbf{Q}$ is calculated in terms of $\mathbf{X}$ via (4.22) and

$$\mathbf{A} \equiv \nabla C(\mathbf{Q}) = \widetilde{\mathbf{H}}^{\dagger}\left[\mathbf{I} + \widetilde{\mathbf{H}}\mathbf{Q}\widetilde{\mathbf{H}}^{\dagger}\right]^{-1}\widetilde{\mathbf{H}}. \tag{4.27}$$

- The perfect gradient feedback assumption can be relaxed to an unbiased noisy estimation that meets the following conditions.

  (H1) *Unbiasedness:*

$$\mathbb{E}\left[\hat{\mathbf{V}}t\mathbf{Q}_{t-1}\right] = 0. \tag{H1}$$

[3] The operator $\mathbf{\Pi}$ denotes the Euclidean matrix projection map: $\mathbf{\Pi}(\mathbf{Y}) = \arg\min_{\mathbf{X} \in \mathscr{X}} \|\mathbf{X} - \mathbf{Y}\|^2$.

Algorithm 4 Onlline Gradient Ascent (OGA)

Require: variable step-size sequence $\gamma_n > 0$

initialize $t \leftarrow 0$, $\mathbf{X}_0 \leftarrow 0$ # initialization

while trasnsmission **do**

$t \leftarrow t+1$;

$\mathbf{Q}(t) \leftarrow \frac{P_c P_{\max}}{P_c + P_{\max}(1-\mathrm{tr}\ \mathbf{X}_{t-1})} \cdot \mathbf{X}_{t-1}$; # pre-transmission phase

transmit;

get $\widetilde{\mathbf{H}}_t$; # post-transmission phase

$\mathbf{A}_t \leftarrow \widetilde{\mathbf{H}}_t^\dagger \left[\mathbf{I} + \widetilde{\mathbf{H}}_t \mathbf{Q}_t \widetilde{\mathbf{H}}_t^\dagger\right]^{-1} \widetilde{\mathbf{H}}_t$;

$\mathbf{V}_t \leftarrow \frac{P_{\max}}{P_c + P_{\max}} \left[\mathbf{A}_t + \frac{\mathrm{tr}(\mathbf{A}_t \mathbf{Q}_t) - C(\mathbf{Q}_t)}{P_c} \cdot \mathbf{I}\right]$;

$\mathbf{X}_t \leftarrow \mathbf{\Pi}\left(\mathbf{X}_{t-1} + \gamma_t \mathbf{V}_t\right)$;

end while

(H2) *Tame error tails:*

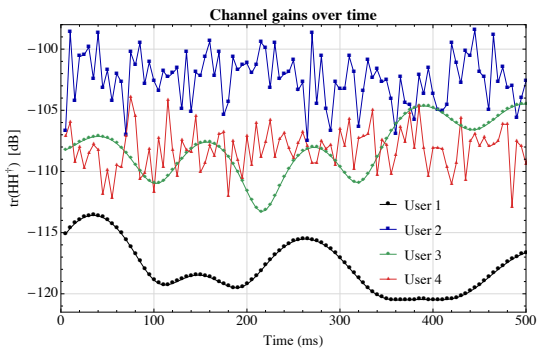
$$\mathbb{P}\left(\|\hat{\mathbf{V}}_t - \mathbf{V}_t\| \geq z\right) \leq B/z^\beta \quad \text{for some } B > 0 \text{ and for some } \beta > 2. \tag{H2}$$

The proposed algorithm is shown to lead to no regret under these conditions.

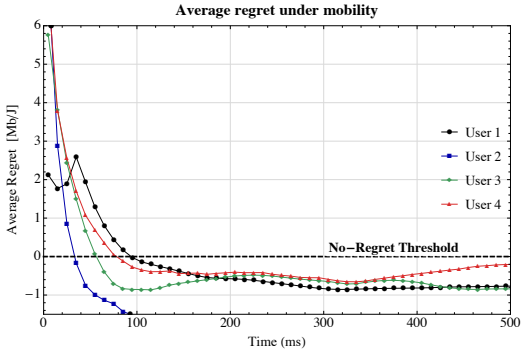
- Our numerical simulations show that, in realistic network environments, users track their individually optimum transmit profile even under rapidly changing channel conditions, achieving gains of up to 600% in energy efficiency over uniform power allocation policies.

To illustrate the performance of our algorithm, we consider in Fig. 4.2 the case of mobile users whose channels vary with time due to Rayleigh fading, path loss fluctuations, etc. For simulation purposes, we used the extended typical urban (ETU) model for the users' environment and the pedestrian (3–5 km/h) and vehicular (30–130 km/h) models in [89]; for reference, the focal users' channel gains ($\mathrm{tr}(\mathbf{H}\mathbf{H}^\dagger)$) have been plotted in Fig. 4.2a. Despite the channels' variability, Fig. 4.2b shows that the users attain a no-regret state in a few iterations, even under rapidly changing channel conditions (cf. the case of Users 2 and 4 with an average speed of 30 km/h and 130 km/h respectively). For completeness, we also plot in Figs. 4.2c and 4.2d the achieved energy efficiency for a pedestrian and a vehicular user, and we compare it to its instantaneous maximum value, the users' initial (uniform) power allocation policy, and the "oracle" solution which corresponds to the best fixed transmit profile in hindsight (i.e. the solution of the offline maximization problem which posits that users can predict the system's evolution in advance). Remarkably, even under rapidly changing channel conditions, the users' achieved energy efficiency tracks its (evolving) maximum value remarkably well and consistently outperforms even the oracle solution (a fact which is consistent with the negative regret observed in Fig. 4.2b).

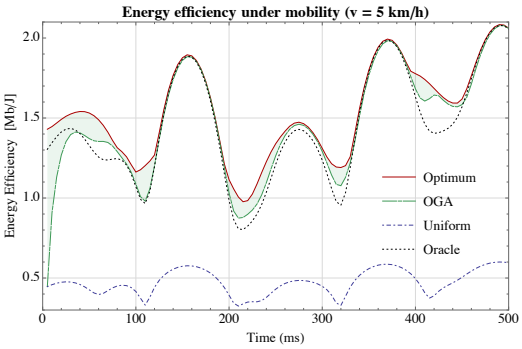
An intuitive explanation for the adaptability of OGA is provided by Figs. 4.2e and 4.2f where we plot the transmit power of the optimum policy, the OGA scheme, and the oracle solution for the same users as in Figs. 4.2c and 4.2d. Even though the optimum covariance matrix $\mathbf{Q}_t^*$ may change significantly from one frame to the next, $\mathrm{tr}(\mathbf{Q}_t^*)$ remains roughly constant (within a few



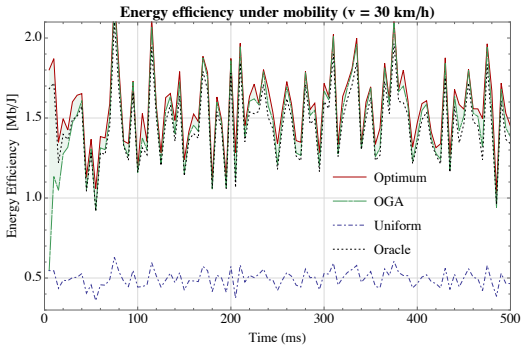
(a) Channel gain evolution for different user velocities



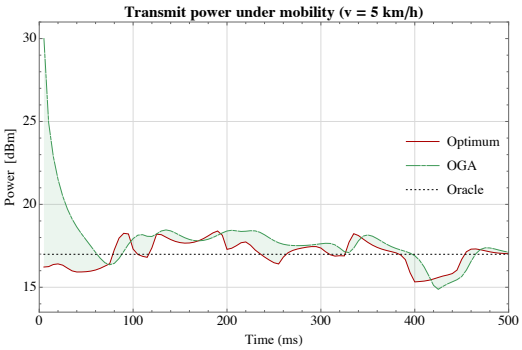
(b) User regret under OGA



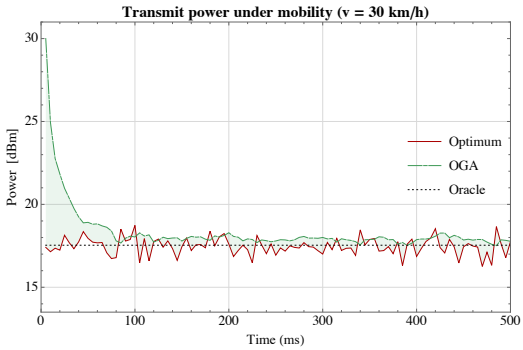
(c) Energy efficiency under OGA (pedestrian)



(d) Energy efficiency under OGA (vehicular)



(e) Transmit power evolution under OGA (pedestrian)



(f) Transmit power evolution under OGA (vehicular)

Figure 4.2: Performance of the OGA algorithm in a dynamic setting with mobile users moving at $v = \{3, 30, 5, 130\}$ km/h. The users' achieved energy efficiency tracks its (evolving) maximum value remarkably well, even under rapidly changing channel conditions.

dBm) over the entire transmission horizon. The OGA algorithm then learns this power level in a few iterations and stays close to it throughout the transmission horizon; as a result, the users' achieved energy efficiency remains itself very close to its maximum value for all time.

We remark that an OMD-based approach could be exploited instead of OGD to maximize the energy efficiency. In so doing, by carefully choosing the regularizer h, the matrix Euclidean projection step could be replaced with a matrix exponential one, which may be easier to compute. Also, better decay rates in terms of the dimensionality of the problem could be obtained. As examples of the matrix exponential learning, the reader is referred to the online rate maximization problem in a cognitive radio network in [J11] or the stochastic optimization framework in [J16].

A nontrivial theoretical question rising in the static (or stochastic) particular case is whether the system converges to an equilibrium state if all users employ a no-regret policy (our numerical simulations show that this indeed the case over a wide region of system parameters in the static case). This question has been investigated in [J16], in which we derive the convergence conditions of the matrix exponential learning algorithm based on OMD for a general class of problems which include the energy-efficiency maximization problems.

Open Issues

An interesting problem is to investigate the regret decay rate of the matrix exponential learning algorithm in [J16] in the general online setting and compare these results with the OGA algorithm. Also a complexity comparaison of the two algorithms could be highly pertinent in the massive MIMO setting, in which the number of users and the number of antennas at the base stations grow large.

Additionally, reducing the matrix-worth of required feedback needed to compute the gradient of the objective at each step is an interesting issue [90]. At last, different and more practical throughput-per-power models could be considered as energy-efficency measures [20], [21].

4.2.2 Feedback-limited IoT Networks

In this section, we overview the work of the PhD student Alexandre Marcastel [C31], [J21], whom I have co-advised (at 40%) jointly with Inbar Fijalkow (ENSEA, 30%, official director) and Panayotis Mertikopoulos (CNRS,30%), funded by the Chair Orange IoT at the University of Cergy-Pontoise Foundation and also supported by ENSEA, Inria and the ANR JCJC ORACLESS.

Regarding the emerging IoT paradigm, which is projected to connect billions of wireless "things" (wireless sensors, wearables, biochip transponders, etc.) in a vast network with drastically different requirements between components (e.g. in terms of throughput and power characteristics) [10], [91], most existing works on resource allocation problems [92], [93], [94] assume that the network remains static over time and the devices are required to have perfect feedback information. In this work, we relax both assumptions by taking into account the inherent dynamics of an IoT network [95], – due itself to the unique mobility attributes of modern wearable devices, intermittent user activity, application diversity etc. – and the impact of feedback imperfections and scarcity.

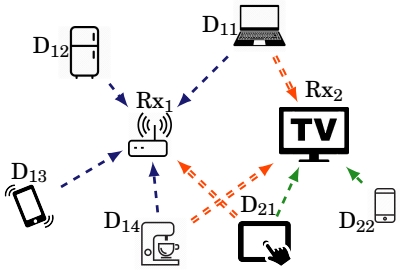


Figure 4.3: System composed of six transmit devices (D_{11}, D_{12}, etc.) and two receivers (Rx_1, Rx_2). The blue and green arrows represent the direct links while the red (double-lined) ones are interfering links.

Problem Formulation

We consider a system composed of M transmitters and N receivers communicating over S orthogonal subcarriers as illustrated in Fig. 4.3: each device transmits to only one intended receiver, but a given receiver may decode several incoming signals. Since we aim at devising a distributed policy that needs no central controller, we can focus on one particular transmitting-receiving pair.

The effective channel gain vector of the focal user $\mathbf{w}_t = (w_t^s)$, where w_t^s represents the effective gain in subcarrier s and is given by

$$w_t^s = \frac{g_t^s}{\sigma^2 + \sum_j g_{j,t}^s p_{j,t}^s}, \quad \forall s, \tag{4.28}$$

where σ^2 is the variance of the noise z_t^s, $p_{j,t}^s$ is the transmitted power by the user j in subcarrier s, $g_{j,t}^s = |h_{j,t}^s|^2$ and $g_t^s = |h_t^s|^2$, and $h_{j,t}^s$, h_t^s denote the channel gains of the focal direct link and tof the interfering link from device j to the focal receiver, respectively.

We consider a power-driven device with a QoS constraint, which aims at minimizing the following loss function:

$$L(\mathbf{p};t) = \sum_{s=1}^{S} p^s \quad + \lambda \left[R_{min} - C(\mathbf{p};t)\right]^+ \tag{4.29}$$

where $\mathbf{p} = (p^1, \ldots, p^S)$ represents the power allocation vector of the focal device with components $p^s, \forall s$ representing the power allocated to the s-th subcarrier. The first term in the objective is the overall power consumption and the second term is a soft-constraint (or penalty) term, which is activated whenever the minimum target rate R_{min} is not achieved. Finally, $C(\mathbf{p};t)$ denotes the well-known Shannon rate[4]:

$$C(\mathbf{p};t) = \sum_{s=1}^{S} \log(1 + w_t^s p^s) \tag{4.30}$$

[4]Regarding the relevance of using this rate measure, which represents the information-theoretic channel capacity of the focal link only under very stringent conditions, the reader is referred to the discussion in Sec. 4.2.1, the second and third paragraphs below equation (4.19).

Algorithm 5 Online exponential learning (OXL)

Require: step-size sequence $\gamma_t > 0$

initialize $t \leftarrow 0$, $\mathbf{y}_0 \leftarrow 0$ # initialization

while trasnsmission **do**

  update $\mathbf{p}_t$ s.t. # pre-transmission phase

  $p_t^s = P_{\max} \frac{\exp(y_t^s)}{1+\sum_{i=1}^{S} \exp(y_t^i)}, \ \forall s;$

  transmit at $\mathbf{p}_t$;

  receive gradient feedback $\mathbf{v}_t$; # post-transmission phase

  $\mathbf{y}_{t+1} \leftarrow \mathbf{y}_t - \gamma_t \mathbf{v}_t$;

  $t \leftarrow t+1$;

end while

and $[x]^+ \triangleq \max\{x, 0\}$, meaning that no penalty is applied when the achieved rate is greater than the threshold $R_t(\mathbf{p}) \geq R_{min}$. Although we choose a linear penalty function for its relevance to communications [C25], [96], [97] (and to simplify the presentation), our results carry over the more general class of concave functions, e.g., logarithmic penalties [98]. The parameter λ can be interpreted as the unit-cost for each bps/Hz under the QoS target R_{min} and it also represents a sensitivity parameter that can be tuned to adjust the flexibility regarding the minimum rate constraint violations or outages.

To sum up, the online optimization problem under study can be stated as:

$$\begin{array}{ll} \text{minimize} & L(\mathbf{p};t) \\ \text{over} & \mathbf{p} = (p^1, \ldots, p^S) \\ \text{subject to} & p^s \geq 0, \ \ \forall s \\ & \sum_{s=1}^{S} p^s \leq P_{\max} \end{array} \tag{4.31}$$

The above *objective function* $L(\mathbf{p};t)$ *may vary in a non-stationary and unpredictable way* such that the focal device cannot determine *a priori* (before the transmission takes place) its instantaneous or dynamic optimal power allocation $\mathbf{p}_t^*$ that minimizes this objective at each time t. Nevertheless, we assume that the device receives some feedback after each transmission, such as the past experienced objective value or its past gradient.

Main Contributions

- We derive an online power allocation policy based on OMD, and which comprises two basic steps: *a*) tracking the gradient (or sub-gradient) of the users' power minimization objective in a dual, unconstrained space; and *b*) using a judiciously designed exponential function to map the output of this step to a feasible power allocation profile and keep going. The proposed procedure is detailed in Algorithm 5.

- To establish a benchmark, we begin with the *full information* or the first-order feedback case, where each wireless device is assumed to have perfect feedback on the gradient of its

individual power minimization objective. In this case, the proposed power allocation policy provided by the OXL is shown in to enjoy the regret guarantee: $\mathrm{Reg}_T/T = \mathcal{O}(T^{-1/2})$ regret guarantee, meaning that the algorithm's performance over a horizon of T transmission cycles is no more than $\mathcal{O}(T^{-1/2})$ away from the best fixed policy in hindsight.

- In addition to providing a comparison baseline, the full information case also allows us to compare the performance of the proposed algorithm to that of classical water-filling algorithms [29], [99], [100] and highlight the difficulties encountered by the latter when the network evolves dynamically over time and only a strictly causal (with no look-ahead) feedback information is available at the transmitter.

- We show that similar regret guarantees can still be attained by OXL even if the feedback received by each device is imperfect and/or otherwise corrupted by non-systematic measurement errors and observational noise. In this case, the received feedback is a gradient estimation denoted by $\tilde{\mathbf{v}}(t)$, which meets the following constraints

$$\begin{aligned} \mathbb{E}[\tilde{\mathbf{v}}_t] &= \nabla L(\mathbf{p}_t), \\ \mathbb{E}[\|\tilde{\mathbf{v}}_t\|^2] &\leq \tilde{V}, \end{aligned} \tag{4.32}$$

  where the expectation is taken over the randomness of the estimator. These conditions are not very restrictive as they require the absence of systematic errors and a bounded variance, as such, they are satisfied by all common error distributions (Gaussian, log-normal, etc) [J16]. For example, the common error model: $\tilde{\mathbf{v}}_t = \nabla L(\mathbf{p}_t; t) + \mathbf{z}$, where $\mathbf{z} \sim \mathcal{N}(0, \sigma_z^2 \mathbf{I})$ [101] satisfies the above conditions.

- On the other hand, if the only information received by each device is the observed value of their past objective function (the so-called *zeroth-order feedback* setting), these bounds change significantly. Lacking any sort of vector-valued, gradient-based feedback, we rely on the simultaneous stochastic approximation technique, which randomly samples the objective function in a neighbourhood of the power policy $\mathbf{p}_t$ to obtain a *(potentially biased) estimate* of the gradient at this point [67], [102].

  The estimator we use in our case is:

$$\tilde{\mathbf{v}}_t = \frac{S}{\delta} L(\tilde{\mathbf{p}}_t; t)\mathbf{u}_t, \tag{4.33}$$

  where $\tilde{\mathbf{p}}_t = \mathbf{p}_t + \delta \mathbf{u}_t$ and $\mathbf{u}_t$ is uniformly taken over the unit Euclidean sphere: $\{\mathbf{u} \in \mathbb{R}^S | \ \|\mathbf{u}_t\|^2 = 1\}$ [67].

  A major problem arises which is that the random sample point $\tilde{\mathbf{p}}_t = \mathbf{p}_t + \delta \mathbf{u}_t$ can fall outside of the feasible set. In our power allocation problem, using the same procedure would imply that the transmit power vector $\tilde{\mathbf{p}}_t$ is allowed to go outside $\mathscr{P}$.

Algorithm 6 Online exponential learning with zeroth-order feedback (OXL_0)

Require: step-size $\gamma > 0$; parameter $0 < \delta \leq P_{\max}/(S+\sqrt{S})$.

initialize $t \leftarrow 0$, $\mathbf{y}_0 \leftarrow 0$ # initialization

while trasnsmission **do**

update $\mathbf{p}_{\delta,t}$ s.t. # pre-transmission phase

$p^s_{\delta,t} = \delta + P_{\max}(1-K_\delta)\frac{\exp(y^s_t)}{1+\sum_{i=1}^S \exp(y^i_t)}, \ \forall s;$

draw a random $\mathbf{u}_t$ uniformly from the unit-sphere

transmit at $\tilde{\mathbf{p}}_t \leftarrow \mathbf{p}_{\delta,t} + \delta \mathbf{u}_t$;

receive scalar feedback $L(\tilde{\mathbf{p}}_t; t)$; # post-transmission phase

compute the gradient estimation $\tilde{\mathbf{v}}_t = \frac{S}{\delta} L(\tilde{\mathbf{p}}_t; t)\, \mathbf{u}_t$

$\mathbf{y}_{t+1} \leftarrow \mathbf{y}_t - \gamma_t \tilde{\mathbf{v}}_t$;

$t \leftarrow t+1$;

end while

One of the main contributions of this work is to introduce a novel learning algorithm that exploits the gradient estimation above, while guaranteeing that the transmit powers always lie in the feasible set. For this, we define a modified and shrunk feasible set $\mathscr{P}_\delta$ such that, for any $\mathbf{p}_{\delta,t} \in \mathscr{P}_\delta$, we have $\mathbf{p}_{\delta,t} + \delta \mathbf{u}_t \in \mathscr{P}$:

$$\mathscr{P}_\delta = \left\{ \mathbf{p}_\delta \in \mathbb{R}^S \;\middle|\; p^s_\delta \geq \delta, \sum_{s=1}^S p^s_\delta \leq P_{\max} - \sqrt{S}\,\delta \right\}. \tag{4.34}$$

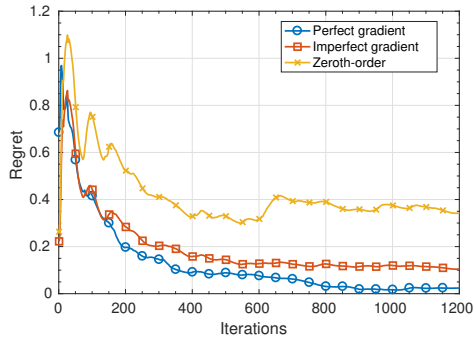
The modified procedure we propose is detailed in Algorithm 6[5].

By jointly optimizing the value of this parameter and that of the original algorithm's step-size, we then show that the proposed policy still leads to no regret at a slower rate: $\frac{\text{Reg}_T}{T} = \mathscr{O}(T^{-3/4})$.

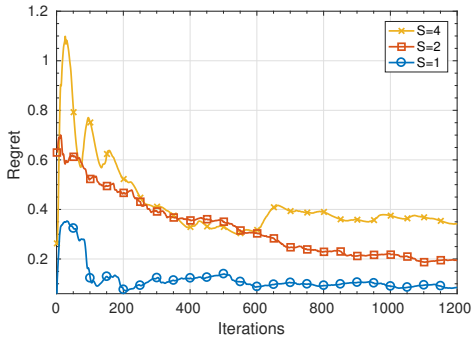
- Therefore, we identify an important tradeoff between the amount of feedback available at the transmitter side and the resulting system performance: if the device has access to unbiased gradient observations, the algorithm's regret after T stages is $\mathscr{O}(T^{-1/2})$ (up to logarithmic factors); on the other hand, if the device only has access to a scalar, utility-based information, this rate drops to $\mathscr{O}(T^{-3/4})$.

- We validate our theoretical analysis via numerical experiments and highlight highly dynamic networks with realistic, unpredictable channel conditions. Classical water-filling algorithms are very sensitive to unpredictable changes in the network and are outperformed by our proposed online algorithms in terms of power consumption and achieved rate. Concerning the impact of available feedback, our numerical results also illustrate a compromise between the amount and/or quality of the feedback information and the algorithms' performance (measured here in terms of the time needed to attain a no-regret

[5]The parameter K_δ in the update is defined by $K_\delta = \frac{\delta}{P_{\max}}(S+\sqrt{S})$.



(a) Regret of OXL and OXL_0 algorithms



(b) Regret of OXL_0 algorithm

Figure 4.4: Impact of feedback amount and problem dimensionality. The average regret of OXL_0 algorithm, relying only on the scalar value of the objective function, decays slower than the average regret of OXL algorithm with perfect or imperfect gradient feedback. Having to estimate the gradient of dimension S using the scalar value of the objective impacts the decay rate of the average regret of OXL_0 algorithm: the higher the problem dimensionality S, the slower the average regret.

state). The zeroth-order feedback case requires only the knowledge of a scalar at each iteration (the value of the objective function) as opposed to a vector (the gradient), but the average time required to reach a no-regret state is higher.

In Fig. 4.4, we illustrate the performance of the OXL_0 algorithm in terms of average regret by investigating the impact of having a scarce or imperfect feedback and the impact of the problem dimensionality S. Fig. 4.4a confirms that having an imperfect gradient feedback does not influence significantly the regret decay rate, as anticipated by our theoretical results. However, this is no longer true when the only information available at the device end is a single scalar. The average regret of the OXL_0 algorithm decays slower compared with OXL algorithm (though the latter cannot be applied with zeroth-order feedback). Finally, Fig. 4.4b illustrates the average regret of OXL_0 algorithm for different values of the problem's dimensionality $S \in \{1,2,4\}$. In all cases, the average regret decays to zero; however, if the number of available subcarriers increases, the variance of the estimator $\tilde{\mathbf{v}}(t)$ increases commensurately. Therefore the quality of the estimator decreases, which results in a reduced decay rate of the average regret.

Open Issues

Interesting extensions for future work are: *i*) reducing even further the required feedback to only one bit of information, which will likely require to adjust the objective functions to be optimized; *ii*) extending the work on zeroth-order feedback to networks composed of multi-antenna devices as mentioned in Sec. 4.2.1, in which the feedback consists of a (potentially large)

matrix; *iii*) exploring efficient means to tune the various parameters of the algorithms for a more practical implementation.

4.2.3 Outputs

We resume below the combined outputs of the work on dynamic and unpredictable wireless networks described in ‘Sec. 4.2.1 (energy-efficient MIMO networks) and Sec. 4.2.2 (feedback-limited IoT networks).

Publications: 2 journals, 4 conf.

[J21] A. Marcastel, **E. V. Belmega**, P. Mertikopoulos, and I. Fijalkow, “Online power optimization in feedback-limited, dynamic and unpredictable IoT networks”, ***accepted paper***, *IEEE Trans. on Signal Processing*, Mar. 2019.

[J16] P. Mertikopoulos, and **E.V. Belmega**, “Learning to be green: Robust energy efficiency maximization in dynamic MIMO-OFDM systems”, *IEEE Journal on Selected Areas in Communication, Special Issue on Energy-Efficient Techniques for 5G Wireless Communication Systems*, vol. 34, no. 4, pp. 743 – 757, Apr. 2016.

[C33] A. Marcastel, **E.V. Belmega**, P. Mertikopoulos, and I. Fijalkow, “Interference Mitigation via Pricing in Time-Varying Cognitive Radio Systems”, **invited paper**, *NetGCoop 2016*, Avignon, France, Nov. 2016.

[C31] A. Marcastel, **E.V. Belmega**, P. Mertikopoulos, and I. Fijalkow, “Online Interference Mitigation via Learning in Dynamic IoT Environments”, *IOE worksop in IEEE GLOBECOM 2016*, Washington DC, USA, 4-8 Dec. 2016.

[C30] A. Marcastel, **E.V. Belmega**, P. Mertikopoulos, and I. Fijalkow, “Online power allocation for opportunistic radio access in dynamic OFDM networks”, *IEEE VTC-Fall 2016*, Montreal, Canada, 18-21 Sep. 2016.

[C27] **E.V. Belmega**, and P. Mertikopoulos, “Learning to be Green: Energy-Efficient Power Allocation in Dynamic Multi-Carrier Systems”, *IEEE VTC-Spring*, Glasgow, Scotland, May 2015.

Collaborators: Panayotis Mertikopoulos (CNRS), Inbar Fijalkow (ENSEA)
Student: Alexandre Marcastel (PhD)
Supported by: Chair Orange IoT, ANR JCJC ORACLESS, Inria, PEPS CNRS-INS2I JCJC Real.net, IoT NEWCOM#, ENSEA

4.3 Conclusions

The online optimization framework coupled with the derived online learning algorithms offers a powerful toolbox to design efficient resource allocation policies in networks that may vary in an arbitrary and unpredictable way. Indeed, no assumptions have been made regarding the underlying dynamics of the network (which can potentially be non-stationary).

The online resource allocation policies we presented here have many desirable properties, such as: a) *distributedness*: users rely on local information and all computations are performed autonomously; b) *statlessness*: users do not need to know the state of the system (e.g., the number of users, the network topology); c) *reinforcement*: users tend to adapt to the network changes and to become more efficient based on their past observations; d) *asynchronicity*: there is no need for a global timer or any signaling between users; e) *rely on strictly causal feedback* information that can be imperfect or scarce; f) *theoretical guarantees* in terms of regret minimization.

The next and final chapter is dedicated to an overall discussion of interesting open issues and prospective research directions.

CHAPTER 5

OPEN ISSUES AND PERSPECTIVES

This HDR thesis is concluded with the description of several open research directions starting with the short- to mid-term perspectives and ending with long-term ones.

5.1 Energy Efficiency Beyond 5G

Below, we discuss several perspectives to improve energy efficiency for the IoT and future generation networks. Indeed, energy efficiency will very likely remain among the top objectives in communications in the years to come: As 5G made significant progress in terms of user experience, low-latency and high data-rates, among the key drivers being envisioned for the 6G are energy harvesting and energy-efficiency technologies in general [103], [104].

5.1.1 Energy-efficient IoT Networks

Our recently funded research project ANR-PRCI ELIOT - *Enabling technologies for IoT* in collaboration with FAPESP, Sao Paulo, Brazil is built around four pillars in terms of objectives for future IoT networks: *i*) energy efficiency; *ii*) security; *iii*) latency; *iv*) self-optimization capabilities. My main research interests relate (but are not limited) to the first and fourth objectives of ELIOT.

The context of ELIOT consists of networks of communicating devices such as: distributed control systems (i.e., autonomous vehicles, delivery drones or other unmanned aerial vehicle (UAV) systems) and sensor networks (i.e., health monitoring, smart cities, smart homes, etc.). These applications depend on reliable and low-energy communication systems, as envisaged by the Internet of Things (IoT) paradigm that will connect billions of wireless “things” (sensors, wearables, biochip transponders, etc.) in a vast network with drastically different characteristics

and requirements in terms of: energy efficiency, security, latency and self-optimization capabilities, and demand for innovative enabling technologies [91], [105], [106], [107], [108].

Because of the unprecedented degree of temporal variability of IoT networks [95] due to the inherent wireless channel characteristics, device mobility attributes, their intermittent activity and behavior, the energy efficiency resource optimization calls for tools from *online optimization and regret minimization*. This allows us to develop new and improved resource allocation algorithms that are capable of adapting on-the-fly to the changing environment and effectively accounting for devices' mobility [C31], [J16], [J21]. The major advantage of such online algorithms is that they do not rely on any assumptions on the dynamics of the underlying network, which, thus, can be completely arbitrary, offering the appropriate flexibility to take into account the specificities of IoT networks.

Non-orthogonal Multiple Access

With billions of interconnected devices, wireless IoT networks are expected to exhibit massive node densities and thus inevitably experience high interference levels. An orthogonal spectrum allocation will likely become unfeasible given the scarcity of RF spectral resources. In most of the work presented in this manuscript (except for Section 3.1.2, [J17]), we have not assumed that the multiple users are scheduled over orthogonal frequency bands; rather, the users can access any of the available frequency bands and suffer from multi-user interference. We have considered single user-decoding is used at the receiver, which treats the multi-user interference as additive noise. Although this decoding technique has the merit of being simple and distributed, the performance of the network is limited by the network interference.

NOMA protocols can be considered instead, which reduce the impact of the multi-user interference by using more complex decoding techniques compared with single-user decoding [109]. NOMA allows overlapping among the signals from different devices by exploiting power- or code-domain multiplexing. In power-domain NOMA, signals from multiple users are superimposed and successive interference cancellation is used at the receiver to decode the messages [110], [111]. Although NOMA can improve the spectrum efficiency, it has several shortcomings that need to be addressed in the IoT context: the multi-user detection increases the decoding complexity and energy consumption of the receivers; the channel state information is often assumed to be perfect; the inherent lack of fairness among the devices.

Energy Harvesting

A possible solution to mitigate the network interference and to prolong the lifetime of IoT devices is energy harvesting from the ambient RF radiation or from dedicated RF-sources (via wireless power transfer) [12]. This has motivated the recent interest surge on RF-based joint wireless information and power transfer [12], [14], [J19], [112], [113], [114] with vast applications to sensor networks, wireless body networks, RFID, etc.

However, there is little work on wireless-powered NOMA networks from an energy-efficiency perspective. Existing works [115], [116], [117], [118] focus on rate maximization objectives with various fairness considerations and assuming the circuit power consumption (to receive/process information) is negligible compared to the power for information transmission. The recent results in [119] question the benefits of NOMA over orthogonal protocols when the circuit power consumption is taken into account. Moreover, none of the cited works seem to take into account the temporal dynamics of the network. All this motivates our interest concerning energy-efficient NOMA transmissions for IoT networks.

One-bit Feedback Information

Aside from having limited processing capabilities, a rising challenge in energy-efficient resource management is the fact that wireless devices in IoT networks receive limited, outdated and/or corrupted feedback from their environment. Therefore, an interesting future work is to reduce even further the amount of information fed back to the devices and to compensate as much as possible for such feedback scarcity. Ideally, we would like to reduce the feedback to a single bit, similarly to the ACK/NACK higher-layer mechanisms.

In Sec. 4.2.2, we have seen that it is possible to do a *gradient descent without a gradient* [102] by using an estimation (possibly biased) of the gradient obtained by stochastic approximation from a single sample of the objective function. In [120], the authors propose a distributed algorithm to maximize the stochastic network utility also based on a zeroth-order feedback and by assuming (full or partial) feedback exchange among the users. An immediate issue to be investigated is whether our online learning algorithms are robust to estimation errors in the value of the objective functions by using the same approach as in the imperfect and unbiased (zero-mean and bounded variance of the estimation errors) gradient feedback case.

A longer-term investigation is to device low-complexity algorithms relying on one-bit of feedback information. For instance, the information bit could be equal to one if $\ell_t(x_t) \geq \rho$ the payoff is above a certain fixed threshold ρ, and zero otherwise. Performing gradient decent based only on this feedback information seems quite challenging. A possible lead and starting point would be to change the problem formulation w.r.t. both its objective function and its feasible set and transform it into a MAB formulation with discrete and finite set of actions [82], [83] with an outage probability QoS constraint instead of an Shannon-rate one [J6], [C10].

5.1.2 Zeroth-order Feedback in Dynamic MIMO Systems

In the same spirit as above, the feedback information in MIMO systems becomes problematic when the number of antennas, users and available frequency subcarriers increases. Classic iterative water-filling algorithms [100] as well as our proposed online learning algorithms in [J16], [J11] (e.g., Sec. 4.2.1) require at each iteration the information regarding the observed gradient, which is a matrix-worth of feedback of dimension dictated by the number of transmit

antennas and the number of available subcarriers per user. This becomes an issue in highly-dense networks, in which such signaling overhead introduce a huge traffic burden [90], [121].

Recently, the authors of [90], [121] have proposed two ways to reducing the amount of feedback information of the exponential matrix algorithm in [J16] by focusing on a static network and investigating the convergence to the Nash equilibrium. The first solution is to receive only sporadically the gradient matrix as opposed to receiving this information at every iteration, while the second alternative assumes that only a part of the entries of the gradient matrix is received at each iteration. In both cases, it is shown the algorithm can still converge to the optimal or Nash equilibrium solution almost surely.

An interesting extension of the work of the PhD student Alexandre Marcastel (ETIS) [J21] described in Sec. 4.2.2 is to propose a zeroth-order feedback online learning algorithm for MIMO systems. The idea is to build an estimator of the gradient matrix relying solely on the scalar value of the objective function and to develop an algorithm that requires a single scalar worth of feedback (instead of the matrix-worth feedback). This extension is a short-term perspective supported by the ANR JCJC ORACLESS project and is ongoing work jointly with the Post-doc Olivier Bilenne (LIG) and Panayotis Mertikopopulos (CNRS).

As also discussed in Section 5.1.1, a further (mid-term) extension would be to reduce the feedback to a single bit of information (quite drastic in MIMO systems) and would likely require to revisit and simplify the problem formulation and to target a less ambitious objective.

5.1.3 More Practical Energy-efficiency and Energy Consumption Models

Investigating more practical energy-efficiency measures and power consumption models is a mid-term perspective that will involve extending the proposed work to non-convex (online) optimization leading to nontrivial resource optimization problems.

As argued in Sec. 4.2.1, the main advantage of investigating the energy efficiency as the tradeoff between the Shannon rate and the power consumption is the simplicity of the resulting optimization problems enabling the devise of simple and efficient algorithms. Most of these problems turn out to be either convex or fractional programs, which can be transformed into equivalent convex problems [3], [19], [122] (see also in Sec. 4.2.1, by exploiting the Charnes-Cooper transformation).

Such energy-efficiency measures have several drawbacks. First, although the Shannon rate is a meaningful information-theoretic measure of the data rate at which reliable information can be transmitted, its simple expression relies often on unrealistic assumptions such as: infinite-length Gaussian codebooks, static channels, perfect channel state information at the receiver, Gaussian distribution not only for the channel noise but also for the interference terms. The last assumption becomes particularly problematic in *arbitrarily dynamic* wireless networks.

This suggests a mid-term perspective by considering the energy efficiency defined in [20], [21], [22] which replaces the Shannon rate with a bit error rate goodput allowing to include

several practical aspects such as: modulation type, access and channel coding techniques, filter architecture, delay and cross-layer aspects, etc. [122], [123]. A second alternative is to study an outage-based goodput introduced in [J6], [C10]. Both possibilities will lead to non-convex optimization problems which do not fall into the fractional programming framework as the objective functions will be ratios of S-shaped functions and linear functions. Nevertheless, the resulting objective is quasi-concave, which can be efficiently optimized in the classical static framework (based on iterative methods [24]). A first idea is to explore the quasi-convex optimization framework and to propose new online learning algorithms suitable to such problems.

A second issue with the considered energy-efficiency measures is the circuit power consumption, which is a critical parameter in the ratio between the Shannon rate (or goodput) and the overall power consumption ratio. Indeed, as shown in [17] and in [J6] for MIMO channels, when the circuit power is ignored entirely, maximizing the energy efficiency leads to the trivial result: the transmitter should remain silent. When the circuit power is a strictly positive constant, the optimal energy-efficient resource allocation policy is highly sensitive to its specific value and, hence, it has to be carefully estimated.

Moreover, it has been shown in [33], [124] that the circuit power consumption is not a fixed constant but depends (either linearly or not) on the number of transmit and receive antennas and on the data rate via the power consumption of the transceiver chains and of the specific processing (channel estimation, channel coding, etc.). Taking all this into account may cast the problem into a non-convex optimization one going beyond the quasi-convex framework above. This challenging issue will require exploring the recent results for non-convex online optimization in [125], which is a long-term objective.

5.2 AI-enabled Communications

One of the most promising set of tools to address the numerous and bold requirements for future communications in terms of: 3D-connectivity, energy efficiency, security, privacy, etc. is artificial intelligence (AI) combining Big Data and machine learning tools. AI-enabled communications, which may be seen as the evolution of cognitive radio (intelligent) networks, have been gaining a lot of momentum [72], [126], [127] and will likely play a predominant role in 6G and future communications [104].

We start with a classic machine learning framework applied to a specific problem (i.e., mmWave beam alignment) and then we move on towards long-term perspectives related to exploring neural networks and *deep learning* for resource management.

5.2.1 Multi-armed Bandits for mmWave Beam Alignment

Because of the congestion of the sub-6 GHz spectrum the millimeter wave (mmWave) band, ranging from 30 GHz to 300 GHz, has been considered as a promising solution for future wireless

networks [128], [129] to achieve the high data rates required by the data-hungry applications. However, propagation at mmWave frequencies is characterized by high pathloss caused by free-space pathloss, penetration loss and absorption by different components of the wireless environment [130], [131]. This suggests exploiting highly directional beams by using large antenna arrays jointly with beamforming techniques [132], [133], [134], [135] to compensate for the high propagation loss. Luckily, the small wavelength of mmWave allows to place a large number of antenna elements in relatively small size arrays which yields a large beamforming gain [136] by focusing the signal's power toward the intended user's equipment.

Recently, machine learning tools have been proposed to solve the beam alignment problem [137], [138]. The authors in [137] proposed an online learning algorithm for mmWave vehicular communications by modeling the beam-alignment problem as a MAB problem with contextual information (i.e., the vehicle's direction of arrival). In [138], the unimodal beam-alignment algorithm exploits the correlation between consecutive beams (as contextual information) and the unimodality of the received power to reduce the search space and to maximize the received energy. The main issue with the proposed algorithms in [137], [138] is that they are centralized policies in that the transmit-beamformer and receive-combiner pair is jointly computed at a central controller and then sent to both the transmitter and receiver, which potentially leads to heavy network signaling.

A promising solution is to adopt an adversarial (as opposed to a stochastic) MAB formulation for the beam-alignment problem that allows us to decouple and split the processing cost between the transmitter and receiver. Both the transmitter and receiver nodes can exploit their own online learning policy to choose their own beam direction in a distributed manner without knowing each others choices. Hence, the learning is carried out independently at both the transmitter and receiver without relying on a central node as in [138] and [137].

This work is currently ongoing jointly with the PhD student Irched Chafaa (ETIS and L2S) co-advised with Mérouane Debbah (Huawei Paris, France) [C38].

5.2.2 Deep Learning for Resource Allocation Problems in Wireless Communications

A long-term perspective is to investigate more advanced machine learning techniques going beyond the MAB and the online optimization framework by exploiting neural networks and *deep learning* [72], [126], [127]. In theory, the promise of such techniques is very ambitious and lies in their capability of learning a completely generic prediction model (or a complex input-output blackbox function), which can be domain-agnostic and, thus, they can - in philosophy - solve any problem! These methods also come with formal and theoretical guarantees via the agnostic *probably approximately correct* (PAC) learning framework [71]. However, in order to achieve such domain-agnostic solutions, these methods require two important ingredients: i) practically limitless training data sets; coupled with ii) practically limitless computational power.

Obviously, such agnostic methods are not efficient in real-life problems, more specifically in wireless communications, and the communications expertise should provide leads on building problem-specific and optimized neural network architectures relying on reasonable amount of training data and computational power. This observation opens vast opportunities of exploiting deep learning for wireless communications, in particular regarding the resolution of resource allocation problems.

Traditional resource allocation problems so far (e.g., energy-efficiency problems such as the ones discussed in Chapter 3 and Chapter 4), often rely on simplified network models and approximations, which have several major advantages: i) they are tractable; ii) they provide non-trivial allocation policies; and iii) they come with strong and (often) meaningful theoretical guarantees. However, when including a lot of practical considerations such as: cross-layer HARQ techniques, channel coding, practical modulations, transceiver architecture, etc., the energy-efficiency maximization problems (but also, rate maximization or resource optimization in general) become very challenging, non-convex and require: either approximation techniques and very specific variable changes and tricks to reduce the initial problem to a tractable proxy; or the use of iterative methods (e.g., alternating optimization techniques) with no or little theoretical guarantees. The reader is referred for instance to [139] for a study of such practical energy-efficient communications. Notice that there is always a non-trivial tradeoff between two extremes: either studying a simple but unpractical model or incorporating all practical considerations leading to an untractable problem.

The deep learning framework provides an alternative solution to the above to solve difficult and non-convex resource allocation problems. In such problems, a first idea would be to start with a tractable simplified model (as in [140], [141]) and gradually incorporate (in a step-by-step manner) practical considerations by using neural networks given their powerful generic modeling properties. This way, each step would correspond to a specific practical aspect that would provide intuitions on how to build a relevant neural network architecture, which would require a reasonable amount of data and computational power. At the opposite, a very large neural network that models all practical considerations at once would be difficult to design, would contain a lot of parameters to be optimized and, thus, would require a large data set and a high computational power.

A first possible critic of such deep learning approaches is that they provide little theoretical guarantees with meaningful insights as opposed to classical approaches and rely mostly on empirical validation. Indeed, the agnostic PAC learning guarantees above remain quite theoretical, they cannot be exploited to design a neural network architecture that is optimized for a specific problem and they generally lead to very loose performance bounds, far from the actual performance of deep learning methods. Nevertheless, when solving practical problems going beyond the convex optimization framework, there are little or no theoretical guarantees anyway and only empirical validation can be performed.

A second critic relates to the size of the available training datasets, which can be a crucial bottleneck depending on the application. For instance, in medical imaging applications the available datasets are very limited [142] due to several issues such as: patient privacy and high cost (as specialized medical doctors have to label the data). In wireless communications as opposed to multimedia applications in general, simulating the data amounts to simulating the multi-user communication channel, which is not very challenging. Indeed, a variety of relevant simulators that have been validated in practical settings are available in the literature, e.g., the COST-HATA model in [143]. Using actual network data may rise security and privacy issues and such data is protected by the network operators for the time being.

At last, regarding the computational power, the most common solution is to use parallel computing techniques (GPGPU or general-purpose processing on graphics processing units), knowing that the most complex steps in the learning procedure is the training of the neural network and is performed offline (only once) using cloud computing, which raises the problem of signaling. Edge-learning at the mobile devices' end is a more challenging issue [72], [95].

Deep neural networks (DNN) are being increasingly exploited for various applications to mobile and wireless networking that range from mobile data analysis, mobility analysis and localization, wireless sensor networks, network control and resource management, security, to signal processing [72]. Regarding resource allocation problems, several recent works [140], [141], [144], [145] show the enormous potential of the deep learning toolbox.

A first approach is to learn the behavior of a complex algorithm using supervised learning and DNN [140], [141]. In [140], the main idea is to treat a given resource optimization algorithm as a black box, and learn its input-output relation by using a deep neural network. The training is done by running the original algorithm on simulated data. In so doing, real-time resource allocation can be performed by simply passing the algorithm input through the network to produce an output, which can be computationally advantageous. This approach is illustrated in a specific problem: the sum-rate maximization problem in an interference network, which is known to be a non-convex problem. Compared with the iterative weighted MMSE algorithm (based on alternate optimization) [146], the DNN is shown to achieve several orders of magnitude speedup. A similar approach is used in [141], in which DNN are trained to solve the sub-band and power allocation problem in a multi-cell network in an aim to maximize the total network throughput (also, a non-convex problem). The training data is obtained by using a genetic algorithm and the results show that the proposed approach provides the optimal solution in 86.3% cases.

Reinforcement deep learning is another approach currently being investigated in resource allocation problems [144], [145]. In [144], the authors exploit DNN to find the Nash equilibrium of a non-cooperative game modeling the resource allocation problem of small base stations in licensed assisted access LTE (LTE-LAA, which overcomes the spectrum scarcity by simultaneously allowing the access to both licensed and unlicensed bands of small base stations). The proposed approach allows the multiple small base stations to perform dynamic channel selection, carrier

aggregation, and fractional spectrum access such that their total throughput is maximized while guaranteeing fairness with existing WiFi networks or other LTE operators in the unlicensed bands. Also, unlike existing methods based on centralized solutions or on coordination among the small base stations, the approach is based on exploiting past-observations of the network and predict the spectrum availability so as to plan ahead the usage of the channels in a distributed way. The proposed deep reinforcement learning scheme is shown to reach the equilibrium, when it converges.

In [145], a dynamic spectrum access in multichannel wireless networks is investigated. The objective is to maximize the expected reward in terms of the achieved rate from both a non-cooperative dynamic game and a centralized perspective. A distributed learning algorithm based on deep learning and reinforcement Q-learning is proposed for dynamic spectrum access that can adapt to real-world network changes while overcoming complexity due to a large state space and partial observability of the problem. Convergence to the solutions (Nash equilibrium or Pareto optimal) is observed in 80% of the experiments.

Rising and interesting challenges include: finding efficient offline training methods to generate optimal or near-optimal sample data, devising efficient (of reduced complexity and possibly problem specific) architectures and configurations of the DNN, investigating DNN in more challenging resource allocation problems as discussed above (including energy-efficiency and other practical considerations), taking into account the dynamics of the wireless networks (potentially non-stationary and arbitrary) and studying *non-equilibrium* online solutions.

CHAPTER 6

BIBLIOGRAPHY

6.1 Our Contributions

[J13] G. Bacci, E. V. Belmega, P. Mertikopoulos, and L. Sanguinetti, "Energy-aware competitive power allocation for heterogeneous networks under QoS constraints.", *IEEE Trans. Wireless Commun.*, vol. 14, no. 9, pp. 4728–4742, 2015.

[J14] E. V. Belmega, L. Sankar, and H. V. Poor, "Enabling data exchange in two-agent interactive systems under privacy constraints", *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 7, pp. 1285–1297, 2015.

[J15] S. Gupta, E. V. Belmega, and M. A. Vázquez-Castro, "Game theoretical analysis of rate adaptation protocols conciliating QoS and QoE", *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, p. 75, 2016.

[C31] A. Marcastel, E. V. Belmega, P. Mertikopoulos, and I. Fijalkow, "Online interference mitigation via learning in dynamic IoT environments", in *IEEE IoE Workshop GLOBECOM 2016*, 2016.

[C30] ——, "Online power allocation for opportunistic radio access in dynamic OFDM networks", in *Vehicular Technology Conference (VTC-Fall), 2016 IEEE 84th*, 2016, pp. 1–5.

[C27] E. V. Belmega and P. Mertikopoulos, "Energy-efficient power allocation in dynamic multi-carrier systems", in *Vehicular Technology Conference (VTC Spring), 2015 IEEE 81st*, 2015, pp. 1–5.

[J16] P. Mertikopoulos and E. V. Belmega, "Learning to be green: Robust energy efficiency maximization in dynamic MIMO–OFDM systems", *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 743–757, 2016.

[J11] ——, "Transmit without regrets: Online optimization in MIMO-OFDM cognitive radio systems", *IEEE J. Sel. Areas Commun.*, vol. 32, no. 11, pp. 1987–1999, 2014.

[J16] P. Mertikopoulos, E. V. Belmega, R. Negrel, and L. Sanguinetti, "Distributed stochastic optimization via matrix exponential learning", *IEEE Trans. Signal Process.*, vol. 65, no. 9, pp. 2277–2290, 2017.

[J8] P. Mertikopoulos, E. V. Belmega, A. L. Moustakas, and S. Lasaulce, "Distributed learning policies for power allocation in multiple access channels", *IEEE J. Sel. Areas Commun.*, vol. 30, no. 1, pp. 96–106, 2012.

[J6] E. V. Belmega and S. Lasaulce, "Energy-efficient precoding for multiple-antenna terminals", *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 329–340, Jan. 2011.

[C10] E. V. Belmega, S. Lasaulce, and M. Debbah, "A survey on energy-efficient communications", in *Proc. IEEE Intl. Symp. Personal, Indoor and Mobile Radio Communications Workshops (PIMRC)*, Istanbul, Turkey, Sep. 2010, pp. 289–294.

[J17] R. Masmoudi, E. V. Belmega, and I. Fijalkow, "Efficient spectrum scheduling and power management for opportunistic users", *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, p. 97, 2016.

[C29] R. Masmoudi, E. V. Belmega, I. Fijalkow, and N. Sellami, "Joint scheduling and power allocation in cognitive radio systems", in *Communication Workshop (ICCW), 2015 IEEE International Conference on*, 2015, pp. 399–404.

[C25] ——, "A unifying view on energy-efficiency metrics in cognitive radio channels", in *Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European*, 2014, pp. 171–175.

[J21] A. Marcastel, E. V. Belmega, P. Mertikopoulos, and I. Fijalkow, "Online power optimization in feedback-limited, dynamic and unpredictable IoT networks", *IEEE Trans. Signal Process.*, 2019, **accepted paper**.

[C18] R. Masmoudi, E. V. Belmega, I. Fijalkow, and N. Sellami, "A closed-form solution to the power minimization problem over two orthogonal frequency bands under QoS and cognitive radio interference constraints", in *Dynamic Spectrum Access Networks (DYSPAN), 2012 IEEE International Symposium on*, 2012, pp. 212–222.

[C23] G. Bacci, E. V. Belmega, and L. Sanguinetti, "Distributed energy-efficient power and subcarrier allocation for OFDMA-based small cells", in *Communications Workshops (ICC), 2014 IEEE International Conference on*, IEEE, 2014, pp. 647–652.

[C22] ——, "Distributed energy-efficient power optimization in cellular relay networks with minimum rate constraints", in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, IEEE, 2014, pp. 7014–7018.

[J20] G. Rezgui, E. V. Belmega, and A. Chorti, "Mitigating jamming attacks using energy harvesting", *IEEE Wireless Commun. Lett.*, 2018.

[J19] E. V. Belmega and A. Chorti, "Protecting secret key generation systems against jamming: Energy harvesting and channel hopping approaches", *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 11, pp. 2611–2626, 2017.

[C34] A. Chorti and E. V. Belmega, "Secret key generation in Rayleigh block fading AWGN channels under jamming attacks", in *Communications (ICC), 2017 IEEE International Conference on*, IEEE, 2017, pp. 1–6.

[C35] E. V. Belmega and A. Chorti, "Energy harvesting in secret key generation systems under jamming attacks", in *Communications (ICC), 2017 IEEE International Conference on*, IEEE, 2017, pp. 1–6.

[C38] I. Chafaa, E. Belmega, and M. Debbah, "Adversarial multi-armed bandit for mmWave beam alignment with one-bit feedback", in *ValueTools 2019, **accepted paper***, 2018.

6.2 Other References

[1] "The 1000x data challenge", Qualcomm, Tech. Rep. [Online]. Available: http://www.qualcomm.com/1000x.

[2] Green Touch Consortium, Tech. Rep. [Online]. Available: http://www.greentouch.org.

[3] A. Zappone, E. Jorswieck, *et al.*, "Energy efficiency in wireless networks via fractional programming theory", *Foundations and Trends® in Communications and Information Theory*, vol. 11, no. 3-4, pp. 185–396, 2015.

[4] Huawei Technologies, *5G: A technology vision*, White paper, 2013.

[5] J. Hoydis, M. Kobayashi, and M. Debbah, "Green small-cell networks", *IEEE Veh. Technol. Mag.*, vol. 6, no. 1, pp. 37–43, 2011, ISSN: 1556-6072.

[6] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?", *IEEE Trans. Wireless Commun.*, vol. 31, no. 2, pp. 160–171, 2013.

[7] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays", *IEEE Signal Process. Mag.*, vol. 30, pp. 40–60, 2013.

[8] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems", *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, 2014.

[9] J. G. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?", *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, 2014.

[10] A. I. Sulyman, S. M. Oteafy, and H. S. Hassanein, "Expanding the cellular-IoT umbrella: An architectural approach", *IEEE Trans. Wireless Commun.*, vol. 24, no. 3, pp. 66–71, 2017.

[11] M. Basharat, W. Ejaz, M. Naeem, A. M. Khattak, and A. Anpalagan, "A survey and taxonomy on nonorthogonal multiple-access schemes for 5G networks", *Transactions on emerging telecommunications technologies*, vol. 29, no. 1, pp. 1–17, 2018.

[12] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless networks with RF energy harvesting: A contemporary survey", *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 757–789, 2015.

[13] R. Ramachandran, V. Sharma, and P. Viswanath, "Capacity of Gaussian channels with energy harvesting and processing cost", *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2563–2575, May 2014.

[14] X. Zhou, R. Zhang, and C. K. Ho, "Wireless information and power transfer in multiuser OFDM systems", *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 2282–2294, 2014.

[15] Y. Gu and S. Aïssa, "RF-based energy harvesting in decode-and-forward relaying systems: Ergodic and outage capacities", *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6425–5434, Nov. 2015.

[16] A. Chorti, K. Papadaki, and H. V. Poor, "Optimal power allocation in block fading channels with confidential messages", *IEEE Trans. Wireless Commun.*, vol. 14, no. 9, pp. 4708–4719, 2015.

[17] S. Verdu, "On channel capacity per unit cost", *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 1019–1030, 1990.

[18] G. Miao, N. Himayat, and G. Y. Li, "Energy-efficient link adaptation in frequency-selective channels", *IEEE Trans. Commun.*, vol. 58, no. 2, 2010.

[19] C. Isheden, Z. Chong, E. Jorswieck, and G. Fettweis, "Framework for link-level energy efficiency optimization with informed transmitter", *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2946–2957, 2012.

[20] D. Goodman and N. Mandayam, "Power control for wireless data", *IEEE Personal Computing*, vol. 7, no. 2, pp. 48–54, 2000.

[21] F. Meshkati, M. Chiang, H. V. Poor, and S. C. Schwartz, "A game-theoretic approach to energy-efficient power control in multicarrier cdma systems", *IEEE J. Sel. Areas Commun.*, vol. 24, no. 6, pp. 1115–1129, 2006.

[22] M. Le Treust and S. Lasaulce, "A repeated game formulation of energy-efficient decentralized power control", *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2860–2869, 2010.

[23] V. Valenta, R. Marsalek, G. Baudoin, M. Villegas, M. Suarez, and F. Robert, "Survey on spectrum utilization in Europe: Measurements, analyses and observations", in *5th International ICST Conference on Cognitive Radio Oriented Wireless Networks and Communications*, 2010, ISBN–978.

[24] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[25] Z. Fei, C. Xing, N. Li, and J. Kuang, "Adaptive multiobjective optimisation for energy efficient interference coordination in multicell networks", *IET Communications*, vol. 8, no. 8, pp. 1374–1383, 2014.

[26] E. Björnson, E. Jorswieck, *et al.*, "Optimal resource allocation in coordinated multi-cell systems", *Foundations and Trends® in Communications and Information Theory*, vol. 9, no. 2–3, pp. 113–381, 2013.

[27] T. Alpcan, T. Başar, R. Srikant, and E. Altman, "CDMA uplink power control as a noncooperative game", *Wireless Networks*, vol. 8, no. 6, pp. 659–670, 2002.

[28] Z. Hasan, G. Bansal, E. Hossain, and V. K. Bhargava, "Energy-efficient power allocation in OFDM-based cognitive radio systems: A risk-return model", *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, 2009.

[29] J.-S. Pang, G. Scutari, F. Facchinei, and C. Wang, "Distributed power allocation with rate constraints in gaussian parallel interference channels", *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3471–3489, 2008.

[30] J. Huang, V. G. Subramanian, R. Agrawal, and R. Berry, "Joint scheduling and resource allocation in uplink OFDM systems for broadband wireless access networks", *IEEE J. Sel. Areas Commun.*, vol. 27, no. 2, pp. 226–234, 2009.

[31] D. T. Ngo and T. Le-Ngoc, "Distributed resource allocation for cognitive radio networks with spectrum-sharing constraints", *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 3436–3449, 2011.

[32] S. Boyd, L. Xiao, and A. Mutapcic, "Subgradient methods", *Lecture Notes of EE392o, Stanford University*, 2003.

[33] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?", *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059–3075, 2015.

[34] S. Hoteit, P. Duhamel, and S. Lasaulce, "On joint power allocation and multipath routing in femto-relay networks", in *Communications (ICC), 2016 IEEE International Conference on*, 2016, pp. 1–6.

[35] A. Pottier, F.-X. Socheleau, and C. Laot, "Robust noncooperative spectrum sharing game in underwater acoustic interference channels", *IEEE J. Ocean. Eng.*, vol. 42, no. 4, pp. 1019–1034, 2017.

[36] G. Scutari, D. Palomar, F. Facchinei, and J.-S. Pang, "Convex optimization, game theory, and variational inequality theory", *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 35–49, Mar. 2010.

[37] G. Miao, N. Himayat, G. Li, and S. Talwar, "Distributed interference-aware energy-efficient power optimization", *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1323–1333, Apr. 2011.

[38] J. Nash, "Non-cooperative games", *Annals of Mathematics*, vol. 54, no. 2, pp. 286–295, 1951.

[39] G. Debreu, "A social equilibrium existence theorem", *Proc. Natl. Acad. Sci. USA*, vol. 38, no. 10, pp. 886–893, Oct. 1952.

[40] F. Facchinei and C. Kanzow, "Generalized Nash equilibrium problems", *Quarterly J. Operations Research*, vol. 5, no. 3, pp. 173–210, Sep. 2007.

[41] D. Fudenberg and J. Tirole, *Game theory*. Cambridge, MA: MIT Press, 1991.

[42] G. Scutari, D. Palomar, and S. Barbarossa, "Optimal linear precoding strategies for wideband noncooperative systems based on game theory part I: Nash equilibria", *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1230–1249, Mar. 2008.

[43] J. G. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?", *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[44] 3GPP Technical Specification Group, "LTE; Requirements for further advancements for Evolved Universal Terrestrial Radio Access (E-UTRA) (LTE-Advanced)", *Tech. Rep. 3GPP TR 36.913 v10.0.0*, Apr. 2011.

[45] C. Ye, S. Mathur, A. Reznik, Y. Shah, W. Trappe, and N. Mandayam, "Information-theoretically secret key generation for fading wireless channels", *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 2, pp. 240–154, Jun. 2010.

[46] T.-H. Chou, S. Draper, and A. M. Sayeed, "Key generation using external source excitation: Capacity, reliability and secrecy exponent", *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2455–2474, Apr. 2012.

[47] W. Yunchuan, Z. Kai, and P. Mohapatra, "Adaptive wireless channel probing for shared key generation based on PID controller", *IEEE Trans. Mobile Comput.*, vol. 12, no. 9, pp. 1842–1852, Sep. 2013.

[48] O. Gungor, F. Chen, and C. Koksal, "Secret key generation via localization and mobility", *IEEE Trans. Veh. Technol.*, vol. 64, no. 6, pp. 2214–2230, Jun. 2015.

[49] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography – Part I: Secret sharing", *IEEE Trans. Inf. Theory*, vol. 39, no. 7, pp. 1121–1132, Jul. 1993.

[50] U. Maurer, "Secret key agreement by public discussion based on common information", *IEEE Trans. Inf. Theory*, vol. 39, no. 5, pp. 733–742, May 1993.

[51] U. Maurer and S. Wolf, "Secret-key agreement over unauthenticated public channels-part III: Privacy amplification", *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 839–851, Apr. 2003.

[52] V. Yakovlev, V. Korzhik, and G. Morales-Luna, "Key distribution protocols based on noisy channels in presence of an acive adversary: Conventional and new versions with parameter optimization", *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2535–2549, Jun. 2008.

[53] S. Mathur, W. Trappe, N. Mandayam, C. Ye, and A. Reznik, "Radio-telepathy: Extracting a cryptographic key from an un-authenticated wireless channel", in *Proc. 14th ACM Annual Int. Conf. Mobile Comput. Netw.*, 2008.

[54] C. Saiki and A. Chorti, "A novel physical layer authenticated encryption protocol exploiting shared randomness", in *IEEE Conf. Commun. Netw. Security (CNS)*, 2015, pp. 113 –118.

[55] M. Zafer, D. Agrawal, and M. Srivatsa, "Limitations of generating a secret key using wireless fading under active adversary", *IEEE/ACM Trans. Netw.*, vol. 20, no. 5, pp. 1440–1451, Oct. 2012.

[56] R. Molière, F. Delaveau, C. L. K. Ngassa, C. Lemenager, T. Mazloum, and A. Sibille, "Tag signals for early authentication and secret key generation in wireless public networks", in *Eur. Conf. Netw. Commun. (EuCNC)*, 2015, pp. 108–112.

[57] X. Song, P. Willett, S. Zhou, and P. Luh, "The MIMO radar and jammer games", *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 687–699, Feb. 2012.

[58] S. Wei, R. Kannan, V. Chakravarthy, and M. Rangaswamy, "CSI usage over parallel fading channels under jamming attacks: A game theory study", *IEEE Trans. Wireless Commun.*, vol. 60, no. 4, pp. 1167–1175, Apr. 2012.

[59] R. El-Bardan, S. Brahma, and P. Varshney, "Strategic power allocation with incomplete information in the presence of a jammer", *IEEE Trans. Commun.*, vol. 64, no. 8, pp. 3467–3479, 2016.

[60] J. Guo, N. Zhao, R. Yu, X. Liu, and V. Leung, "Exploiting adversarial jamming signals for energy harvesting in interference networks", *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 1267 –1280, 2017.

[61] H. Fang, L. Xu, and K. Choo, "Stackelberg game based relay selection for physical layer security and energy efficiency enhancement in cognitive radio networks", *Applied Mathematics and Computation*, vol. 296, pp. 153–167, 2017.

[62] A. Mukherjee and A. Swindlehurst, "Jamming games in the MIMO wiretap channel with an active eavesdropper", *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 82–91, Jan. 2013.

[63] M. Bloch and J. Barros, *Physical-layer security: From information theory to security engineering*. Cambridge, UK: Cambridge University Press, 2011.

[64] S. Ulukus, A. Yener, E. Erkip, O. Simeone, M. Zorzi, P. Grover, and K. Huang, "Energy harvesting wireless communications: A review of recent advances", *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 360–381, Mar. 2015.

[65] R. Zhang and C. K. Ho, "MIMO broadcasting for silmutaneous wireless information and power transfer", *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May 2013.

[66] G. Amariucai, S. Wei, and R. Kannan, "Gaussian jamming in block-fading channels under long term power constraints", in *Proc. Int. Symp. Inf. Theory (ISIT)*, IEEE, Nice, France, 2007, pp. 1001–1005.

[67] S. Shalev-Shwartz, "Online learning and online convex optimization", *Foundations and trends in machine learning*, vol. 4, no. 2, pp. 107–194, 2011.

[68] S. Bubeck and N. Cesa-Bianchi, "Regret analysis of stochastic and nonstochastic multi-armed bandit problems", *Foundations and trends in machine learning*, vol. 5, no. 1, pp. 1–122, 2012.

[69] N. Cesa-Bianchi and G. Lugosi, *Prediction, learning, and games*. Cambridge University Press, 2006.

[70] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An introduction to statistical learning*. Springer, 2013, vol. 112.

[71] S. Shalev-Shwartz and S. Ben-David, *Understanding machine learning: From theory to algorithms*. Cambridge University Press, 2014.

[72] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey", *Arxiv preprint arxiv:1803.04311*, 2018.

[73] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks", *IEEE Trans. Wireless Commun.*, vol. 24, no. 2, pp. 98–105, 2017.

[74] J. Hannan, "Approximation to Bayes risk in repeated play", in *Contributions to the theory of games, volume III*, ser. Annals of Mathematics Studies, M. Dresher, A. W. Tucker, and P. Wolfe, Eds., vol. 39, Princeton, NJ: Princeton University Press, 1957, pp. 97–139.

[75] M. Zinkevich, "Online convex programming and generalized infinitesimal gradient ascent", in *Icml '03: Proceedings of the 20th international conference on machine learning*, 2003, pp. 928–936.

[76] E. Hazan, A. Agarwal, and S. Kale, "Logarithmic regret algorithms for online convex optimization", *Machine learning*, vol. 69, no. 2-3, pp. 169–192, 2007.

[77] A. S. Nemirovski and D. B. Yudin, *Problem complexity and method efficiency in optimization*. New York, NY: Wiley, 1983.

[78] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire, "Gambling in a rigged casino: The adversarial multi-armed bandit problem", in *IEEE Symposium on Foundations of Computer Science (FOCS)*, IEEE, 1995, p. 322.

[79] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online learning for matrix factorization and sparse coding", *Journal of machine learning research*, vol. 11, pp. 19–60, 2010.

[80] D. Garber and E. Hazan, "Adaptive universal linear filtering", *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1595–1604, 2013.

[81] O. Shamir and S. Shalev-Shwartz, "Matrix completion with the trace norm: Learning, bounding, and transducing", *Journal of machine learning research*, vol. 15, pp. 3401–3423, 2014.

[82] A. Anandkumar, N. Michael, A. K. Tang, and A. Swami, "Distributed algorithms for learning and cognitive medium access with logarithmic regret", *IEEE J. Sel. Areas Commun.*, vol. 29, no. 4, pp. 731–745, 2011.

[83] M. Hashemi, A. Sabharwal, C. E. Koksal, and N. B. Shroff, "Efficient beam alignment in millimeter wave systems using contextual bandits", *Arxiv preprint arxiv:1712.00702*, 2017.

[84] J. Wang, C. Jiang, H. Zhang, X. Zhang, V. C. Leung, and L. Hanzo, "Learning-aided network association for hybrid indoor LiFi-WiFi systems", *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3561–3574, 2018.

[85] J. Kwon and P. Mertikopoulos, "A continuous-time approach to online optimization", *Journal of dynamics and games*, vol. 4, no. 2, pp. 125–148, 2017.

[86] X. Ge, X. Huang, Y. Wang, M. Chen, Q. Li, T. Han, and C.-X. Wang, "Energy efficiency optimization for MIMO-OFDM mobile multimedia communication systems with qos constraintsmobile multimedia communication systems with QoS constraints", *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2127–2138, 2014.

[87] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks", *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1089–1098, 2004.

[88] A. Charnes and W. W. Cooper, "Programming with linear fractional functionals", *Naval Research Logistics Quarterly*, vol. 9, pp. 181–196, 1962.

[89] 3GPP, *User equipment (UE) radio transmission and reception*, White paper, Technical Specification, Jun. 2014.

[90] W. Li, M. Assaad, G. Ayache, and M. Larranaga, "Matrix exponential learning for resource allocation with low informational exchange", in *IEEE 19th International Workshop on Signal Processing Advances in*

Wireless Communications (SPAWC), 2018, pp. 266–270. [Online]. Available: https://arxiv.org/pdf/1802.06652v2.pdf.

[91] C. Goursaud and J.-M. Gorce, "Dedicated networks for IoT: PHY/MAC state of the art and challenges", *EAI endorsed transactions on Internet of Things*, 2015.

[92] H. Safdar, N. Fisal, R. Ullah, W. Maqbool, F. Asraf, Z. Khalid, and A. Khan, "Resource allocation for uplink M2M communication: A game theory approach", in *Wireless Technology and Applications (ISWTA), 2013 IEEE Symp.*, (Kuching, Malaysia), 2013, pp. 48–52.

[93] M. S. Ali, H. Tabassum, and E. Hossain, "Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (NOMA) systems", *IEEE Access*, vol. 4, pp. 6325–6343, 2016.

[94] T. Zheng, Y. Qin, H. Zhang, and S. Kuo, "Adaptive power control for mutual interference avoidance in industrial Internet-of-Things", *China Communications*, vol. 13, no. Supplement 1, pp. 124–131, 2016.

[95] T. Chen, S. Barbarossa, X. Wang, G. B. Giannakis, and Z.-L. Zhang, "Learning and management for Internet-of-Things: Accounting for adaptivity and scalability", *Arxiv preprint arxiv:1810.11613*, 2018.

[96] T. Alpcan, T. Başar, R. Srikant, and E. Altman, "CDMA uplink power control as a noncooperative game", *Wireless Networks*, vol. 8, no. 6, pp. 659–670, 2002.

[97] E. Altman and L. Wynter, "Equilibrium, games, and pricing in transportation and telecommunication networks", *Networks and Spatial Economics*, vol. 4, no. 1, pp. 7–21, 2004.

[98] M. Chiang, P. Hande, T. Lan, C. W. Tan, *et al.*, "Power control in wireless cellular networks", *Foundations and Trends in Networking*, vol. 2, no. 4, pp. 381–533, 2008.

[99] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels", *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 145–152, 2004.

[100] G. Scutari, D. P. Palomar, and S. Barbarossa, "The MIMO iterative waterfilling algorithm", *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1917–1935, 2009.

[101] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge University Press, 2005.

[102] A. D. Flaxman, A. T. Kalai, and H. B. McMahan, "Online convex optimization in the bandit setting: Gradient descent without a gradient", in *SODA'05: Proceedings of the 16th annual ACM-SIAM symposium on discrete algorithms*, (Vancouver, British Columbia), 2005, pp. 385–394.

[103] H. Haas, X. You, J. Elmirghani, and K. David, *6G: What is Next?*, IEEE Vehicular Technology Magazine Special Issue, Call for papers, 2018.

[104] E. C. Strinati, S. Barbarossa, J. L. Gonzalez-Jimenez, D. Kténas, N. Cassiau, and C. Dehos, "6G: The next frontier", *Arxiv preprint arxiv:1901.03239*, 2019.

[105] L. Wang, C. Goursaud, N. Nikaein, L. Cottatellucci, and J.-M. Gorce, "Cooperative scheduling for coexisting body area networks", *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 123–133, 2013.

[106] D. Singh, G. Tripathi, and A. J. Jara, "A survey of Internet-of-Things: Future vision, architecture, challenges and services", in *Internet of things (WF-IoT), 2014 IEEE World Forum on*, 2014, pp. 287–292.

[107] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: A survey on enabling technologies, protocols, and applications", *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, 2015.

[108] A. I. Sulyman, S. M. A. Oteafy, and H. S. Hassanein, "Expanding the cellular-IoT umbrella: An architectural approach", *IEEE Trans. Wireless Commun.*, vol. 24, no. 3, pp. 66–71, 2017.

[109] M. Basharat, W. Ejaz, M. Naeem, A. M. Khattak, and A. Anpalagan, "A survey and taxonomy on nonorthogonal multiple-access schemes for 5G networks", *Transactions on Emerging Telecommunications Technologies*, vol. 29, no. 1, 2018.

[110] S. R. Islam, N. Avazov, O. A. Dobre, and K.-S. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges", *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 721–742, 2017.

[111] M. Shirvanimoghaddam, M. Dohler, and S. J. Johnson, "Massive non-orthogonal multiple access for cellular IoT: Potentials and limitations", *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 55–61, 2017.

[112] Z. Ding, S. M. Perlaza, I. Esnaola, and H. V. Poor, "Power allocation strategies in energy harvesting wireless cooperative networks", *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 846–860, 2014.

[113] X. Lu, P. Wang, D. Niyato, and Z. Han, "Resource allocation in wireless networks with RF energy harvesting and transfer", *IEEE Network*, vol. 29, no. 6, pp. 68–75, 2015.

[114] S. B. Amor, S. M. Perlaza, I. Krikidis, and H. V. Poor, "Feedback enhances simultaneous energy and information transmission in multiple access channels", in *Information Theory (ISIT), 2016 IEEE International Symposium on*, IEEE, 2016, pp. 1974–1978.

[115] Y. Yuan and Z. Ding, "The application of non-orthogonal multiple access in wireless powered communication networks", in *Signal Processing Advances in Wireless Communications (SPAWC), 2016 IEEE 17th International Workshop on*, 2016, pp. 1–5.

[116] P. D. Diamantoulakis, K. N. Pappi, Z. Ding, and G. K. Karagiannidis, "Wireless-powered communications with non-orthogonal multiple access.", *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8422–8436, 2016.

[117] H. Chingoska, Z. Hadzi-Velkov, I. Nikoloska, and N. Zlatanov, "Resource allocation in wireless powered communication networks with non-orthogonal multiple access", *IEEE Commun. Lett.*, vol. 5, no. 6, pp. 684–687, 2016.

[118] P. D. Diamantoulakis, K. N. Pappi, G. K. Karagiannidis, H. Xing, and A. Nallanathan, "Joint downlink/uplink design for wireless powered networks with interference", *IEEE Access*, vol. 5, pp. 1534–1547, 2017.

[119] Q. Wu, W. Chen, D. W. K. Ng, and R. Schober, "Spectral and energy efficient wireless powered IoT networks: NOMA or TDMA?", *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 6663 –6667, 2018.

[120] W. Li, M. Assaad, and P. Duhamel, "Distributed stochastic optimization in networks with low informational exchange", in *Communication, Control, and Computing (Allerton), 2017 55th Annual Allerton Conference on*, 2017, pp. 1160–1167.

[121] W. Li and M. Assaad, "Matrix exponential learning schemes with low informational exchange", *Arxiv preprint arxiv:1802.06652*, 2018. [Online]. Available: https://arxiv.org/pdf/1802.06652.pdf.

[122] X. Leturc, C. J. Le Martret, and P. Ciblat, "Energy efficient resource allocation for HARQ with statistical CSI in multiuser ad hoc networks", in *Communications (ICC), 2017 IEEE International Conference on*, IEEE, 2017, pp. 1–6.

[123] M. Maaz, P. Mary, and M. Hélard, "Energy minimization in HARQ-I relay-assisted networks with delay-limited users", *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 6887–6898, 2017.

[124] Björnson, Emil and Hoydis, Jakob and Sanguinetti, Luca, "Massive MIMO networks: Spectral, energy, and hardware efficiency", *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.

[125] E. Hazan, K. Singh, and C. Zhang, "Efficient regret minimization in non-convex games", in *Proceedings of the 34th International Conference on Machine Learning*, vol. 70, International Convention Centre, Sydney, Australia, 2017, pp. 1433–1441. [Online]. Available: http://proceedings.mlr.press/v70/hazan17a.html.

[126] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer", *IEEE Trans. on Cogn. Commun. Netw.*, vol. 3, no. 4, pp. 563–575, 2017.

[127] C. Lee, H. B. Yilmaz, C.-B. Chae, N. Farsad, and A. Goldsmith, "Machine learning based channel modeling for molecular MIMO communications", in *Signal Processing Advances in Wireless Communications (SPAWC), 2017 IEEE 18th International Workshop on*, IEEE, 2017, pp. 1–5.

[128] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!", *IEEE Access*, vol. 1, pp. 335–349, 2013.

[129] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave communications (mmWave) for 5G: Opportunities and challenges", *Wireless Networks*, vol. 21, no. 8, pp. 2657–2676, 2015.

[130] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang, "Overview of millimeter wave communications for fifth-generation 5G wireless networks with a focus on propagation models", *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6213–6230, 2017.

[131] I. Hemadeh, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, "Millimeter-wave communications: Physical channel models, design considerations, antenna constructions and link-budget", *IEEE Commun. Surveys Tuts.*, 2017.

[132] S. Haghighatshoar and G. Caire, "The beam alignment problem in mmWave wireless networks", in *Signals, Systems and Computers, 2016 50th Asilomar Conference on*, IEEE, 2016, pp. 741–745.

[133] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems", *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, 2016.

[134] A. F. Molisch, V. V. Ratnam, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda, "Hybrid beamforming for massive MIMO: A survey", *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 134–141, 2017.

[135] X. Song, S. Haghighatshoar, and G. Caire, "A scalable and statistically robust beam alignment technique for mm-Wave systems", *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4792 –4805, 2018.

[136] ——, "Efficient beam alignment for mmWave single-carrier systems with hybrid MIMO transceivers", *Arxiv preprint arxiv:1806.06425*, 2018.

[137] A. Asadi, S. Müller, G. H. A. Sim, A. Klein, and M. Hollick, "FML: Fast Machine Learning for 5G mmWave Vehicular Communications", in *2018 IEEE International Conference on Computer Communications (INFOCOM)*, 2018.

[138] M. Hashemi, A. Sabharwal, C. E. Koksal, and N. B. Shroff, "Efficient beam alignment in millimeter wave systems using contextual bandits", in *2018 IEEE International Conference on Computer Communications (INFOCOM)*, 2018, pp. 2393 –2401.

[139] X. Leturc, "Allocation de ressources pour les HARQ dans les réseaux ad hoc mobiles", PhD thesis, PhD thesis, Télécom ParisTech., 2018.

[140] H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos, "Learning to optimize: Training deep neural networks for wireless resource management", in *Signal Processing Advances in Wireless Communications (SPAWC), 2017 IEEE 18th International Workshop on*, 2017, pp. 1–6. [Online]. Available: https://arxiv.org/pdf/1705.09412.pdf.

[141] K. Ahmed, H Tabassum, and E Hossain, "Deep learning for radio resource allocation in multi-cell networks", *Arxiv preprint arxiv:1808.00667*, 2018.

[142] J. Bernal, A. Histace, M. Masana, Q. Angermann, C. Sánchez-Montes, C. R. de Miguel, M. Hammami, A. García-Rodríguez, H. Córdova, and O. Romain, "GTCreator: A flexible annotation tool for image-based datasets", *International Journal of Computer Assisted Radiology and Surgery*, pp. 1–11, 2018.

[143] G. F. Pedersen, *COST 231-Digital mobile radio towards future generation systems*. EU, 1999.

[144] U. Challita, L. Dong, and W. Saad, "Proactive resource management for LTE in unlicensed spectrum: A deep learning perspective", *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4674 –4689, 2018.

[145] O. Naparstek and K. Cohen, "Deep multi-user reinforcement learning for distributed dynamic spectrum access", *IEEE Trans. Wireless Commun.*, no. 1, pp. 310 –323, 2018.

[146] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel", *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, 2011.

APPENDIX A

APPENDIX A: SELECTED PUBLICATIONS

This appendix contains five publications among the most significant ones related to energy-efficient wireless communications. The order in which they are listed is relevant to the structure of the manuscript.

[J17] R. Masmoudi, **E.V. Belmega**, and I. Fijalkow, "Efficient Spectrum Scheduling and Power Management for Opportunistic Users", *EURASIP Journal on Wireless Communications and Networking (JWCN)*, vol. 2016:97, pp. 1 – 19, Apr. 2016.

[J13] G. Bacci, **E.V. Belmega**, P. Mertikopoulos, and L. Sanguinetti, "Energy-Aware Competitive Power Allocation for Heterogeneous Networks Under QoS Constraint", *IEEE Trans. on Wireless Communications*, vol. 14, no. 9, pp. 4728 – 4742, Sep. 2015.

[J19] **E.V. Belmega**, and A. Chorti, "Protecting secret key generation systems against jamming: Energy harvesting and channel hopping approaches", *IEEE Trans. on Information Forensics & Security*, vol. 12, no. 11, pp. 2611 – 2626, Nov. 2017.

[J16] P. Mertikopoulos, and **E.V. Belmega**, "Learning to be green: robust energy efficiency maximization in dynamic MIMO-OFDM systems", *IEEE Journal on Selected Areas in Communication, Special Issue on Energy-Efficient Techniques for 5G Wireless Communication Systems*, vol. 34, no. 4, pp. 743 – 757, Apr. 2016.

[J21] A. Marcastel, **E.V. Belmega**, P. Mertikopoulos, and I. Fijalkow, "Online power optimization in feedback-limited, dynamic and unpredictable IoT networks", ***accepted paper***, *IEEE Trans. on Signal Processing*, Mar. 2019.

Masmoudi *et al. EURASIP Journal on Wireless Communications and Networking* (2016) 2016:97
DOI 10.1186/s13638-016-0594-4

EURASIP Journal on Wireless Communications and Networking

RESEARCH Open Access

Efficient spectrum scheduling and power management for opportunistic users

Raouia Masmoudi[1,2*], E. Veronica Belmega[1,3] and Inbar Fijalkow[1]

Abstract

In this paper, we study the centralized spectrum access and power management for several opportunistic users, secondary users (SUs), without hurting the primary users (PUs). The radio resource manager's objective is to minimize the overall power consumption of the opportunistic system over several orthogonal frequency bands under constraints on the minimum quality of service (QoS) and maximum peak and average interference to the PUs. Given the opposing nature of these constraints, we first study the problem of feasibility, and we provide sufficient conditions and necessary conditions for the existence of a solution. The main challenge lies in the non-convexity of this problem because of the discrete spectrum scheduling: one band can be allocated to at most one SU to avoid interference impairments. To overcome this issue, we use a Lagrangian relaxation technique, and we prove that the discrete solutions of the relaxed problem are the solutions to the initial problem. We propose a projected sub-gradient algorithm to compute the solution, when it exists. Assuming that the channels are drawn randomly from a continuous distribution, this algorithm converges to the optimal solution. We also study a specific symmetric system for which we provide the analytical solution. Our numerical results compare the energy-efficiency of the proposed algorithm with other spectrum allocation solutions and show the optimality of our approach.

Keywords: Cognitive radio systems, Spectrum scheduling, Power allocation, Lagrangian relaxation, Projected sub-gradient algorithm

1 Introduction

Most frequency bands in the radio spectrum have already been licensed, and it is difficult to find vacant bands for wireless communication systems. At the same time, the most allocated spectrum is under-utilized [1]. Cognitive radio (CR) systems, as explained in [2] and references therein, propose to better utilize the spectrum by allowing an opportunistic access to it. A hierarchy between users is imposed, in which secondary users (SUs) are allowed, by the spectrum manager, to communicate either in the vacant bands left by the licensed users, called primary users (PUs), or the non-vacant bands under the condition that the created interference (at the primary receivers) is kept below some predefined thresholds [3]. The radio resource manager uses channel state information (CSI) to coordinate the access to the wireless radio spectrum. When performed in a centralized way, this management is often referred to as coordinated multi-point (CoMP) radio resource management [4]. In the CR paradigm, CSI is provided by spectrum sensing at different remote locations and/or by backhaul information feedback from the spectrum manager to improve the spectrum usage. Carrier aggregation and multi-carrier communications have been suggested as promising candidates for both the CR and CoMP systems thanks to their flexible usage of the spectrum [5].

In ad-hoc and sensor networks or even in future 5G, a major bottleneck is the power consumption efficiency caused by limited battery-life device systems and operating costs [6, 7]. In this work, we investigate a centralized power minimization problem with quality of service (QoS) requirements for the secondary users imposed by the spectrum manager. In such a centralized setting, the spectrum manager can be more effective when it operates opportunistic users scheduling in addition to the sole power allocation, as mentioned in [8] and [9]. Very few existing works consider both bandwidth scheduling

*Correspondence: raouia.masmoudi@ensea.fr
[1]ETIS/ ENSEA, University Cergy Pontoise, CNRS, F-95000 Cergy Pontoise, France
[2]LETI, ENIS and University of Sfax, Sfax, Tunisia
Full list of author information is available at the end of the article

and power allocation jointly in the CR context. In particular, [10] provides a heuristic algorithm for the users scheduling.

In this paper, we consider a joint discrete scheduling and power allocation problem that aims at a minimal power consumption under QoS and interference power constraints in a centralized CR system. Such a problem introduces two major challenges. First, the minimum QoS and maximum interference constraints may not be simultaneously satisfied. Several works in the wireless communications literature [11, 12] have proposed a classical water-filling procedure to solve rate-driven or power-driven resource allocation problems in classical interference or multiple access channels. In our study, the presence of the PUs imposes additional peak and total interference constraints aside the classical minimum QoS constraint for the SU communication. These additional constraints are the main reason why classical numerical approaches are not suitable. In order to tackle the problem of feasibility, we study necessary conditions and sufficient conditions on the CSI for the existence of a feasible allocation point. The second difficulty raised by the problem under investigation is the discrete nature of channel assignments in the scheduling policy. This policy makes the problem a non-convex optimization one. Inspired by the approach in [13], we use a Lagrangian relaxation and a dual approach to obtain a solvable convex optimization problem. We then study the Karush-Kuhn-Tucker (KKT) conditions of the relaxed problem and show that the solutions meeting these conditions are actually the solutions of the initial non-convex problem. To solve the relaxed problem numerically, we propose a projected sub-gradient algorithm [14] when the problem is feasible.

1.1 Related works

Power allocation problems have been the subject of several studies in non-CR systems from a rate maximization point of view and via centralized [15] or decentralized (using non-cooperative games) [16–18] approaches. Also, power allocation problems without spectrum allocation in non-CR networks have been studied from an energy-efficiency point of view in centralized [19, 20] and decentralized systems [21, 22]. In this paper, we consider a centralized radio management, to make the spectrum manager more effective, it operates opportunistic user's scheduling in addition to the sole power allocation, as mentioned in [8] and [9]. Such joint resource allocation problems have been the subject of several studies in non-cognitive radio settings such as in code division multiple access (CDMA) systems [23] and in downlink and uplink orthogonal frequency division multiplexing (OFDM) systems [24] and [13], respectively. In [23], the scheduling and resource allocation problem for the downlink in a CDMA-based wireless network is considered. The problem is to select a subset of the users for transmission and, for each of these users, to choose the optimal modulation, coding scheme and transmit power allocation policy. In [24], the authors consider the scheduling and resource allocation for the downlink of a cellular OFDM system. An optimal algorithm is proposed assuming that multiple users can time-share each tone and several low complexity heuristics are that enforce integer tone allocations. Among the works on OFDM [13, 24], the closest to our work is [13], in which a dynamic scheduling and power allocation algorithm was proposed to compute the policies of the multiple non-interfering users that maximize the overall QoS. An algorithm is derived (without the interference constraints of the cognitive radio context) using a Lagrangian relaxation technique to overcome the discrete scheduling constraints. However, a rigorous proof of the convergence and optimality of the proposed algorithm is not provided.

In the cognitive radio context, in which additional interference constraints to protect the primary users must be taken into account, the rate maximization problem was studied in [3, 6, 11]. In these works, the authors consider decentralized solutions in MIMO systems via non-cooperative game theory without spectrum scheduling constraints. The major disadvantage of such decentralized approaches is that the Nash equilibrium solution (i.e., the natural solution concept in non-cooperative games) provides an operating point that is often outperformed by a centralized solution. Other works study rate maximization problems under different CSI assumptions. In particular, [25] addresses the scheduling aspect with partial-CSI at the SU which limits the adaptability to the actual channel state. For a time-varying system, the authors of [26] study dynamic cognitive radio settings without explicit interference temperature constraints imposed by the primary users' presence. Online optimization and no-regret distributed learning algorithms are used in [26] assuming the users do not know the perfect CSI prior to their transmissions. Such a complex approach is not required here, as we consider that CSI is available to the centralized system manager.

Energy efficiency problems in the CR context were studied with QoS and spectrum scheduling constraints in [10], in which the authors minimize the SU's power consumption. The framework in [10] is the closest among all cited references to our paper. However, there is no proof of optimality of the proposed scheduling, which is based on a heuristic method involving some exhaustive search steps. In our work, we use convex optimization tools to find the optimal joint scheduling and power allocation under interferences and QoS constraints. Our optimal solution is calculated via an iterative sub-gradient algorithm that is proven to converge to the optimal solution.

1.2 Our contributions

The main contributions of this paper are summarized here below:

- We derive necessary conditions and sufficient conditions for the existence of a solution to the joint spectrum scheduling and power allocation problem in a CR system.
- We introduce a convex optimization problem based on Lagrangian relaxation of the initial non-convex problem. Then, we prove that the discrete solutions of the relaxed problem are the solutions to the initial problem.
- The optimal solution of the relaxed problem, when it exists, is computed via a projected sub-gradient algorithm. We prove that, when the problem is feasible, our proposed projected sub-gradient algorithm converges to an optimal solution that satisfies the KKT conditions.
- We also study the specific case of a symmetric system for which our iterative algorithm is not suitable and we solve it analytically.
- Numerical results illustrate the energy-efficiency of the proposed allocation strategy compared with other spectrum allocation policies.

The remainder of this paper is organized as follows. The system model is presented in Section 2. In Section 3, we study the joint scheduling and power allocation problem by discussing its feasibility, the Lagrangian relaxation, and its optimality, and then provide a sub-gradient algorithm to solve this problem. Selected numerical results are illustrated in Section 4 to show the efficiency of the proposed solution. Particular cases, for which an analytical solution is found, are studied in Section 5. Finally, Section 6 concludes the paper.

2 System model

We focus on the CR model in Fig. 1 composed of $Q \geq 1$ SUs and $K \geq 1$ PUs. Each primary/secondary user consists of a primary/secondary transmitter (PT/ST) and a primary/secondary receiver (PR/SR), respectively. The transmission is performed over N orthogonal frequency bands. The transmit power of ST_q (of the q^{th} SU) in frequency band $n \in \mathcal{N} \triangleq \{1,\ldots,N\}$ is denoted by p_{qn}, the power allocation of the q^{th} SU is denoted by $\underline{p}_q = \left(p_{q1}, p_{q2}, \ldots, p_{qN}\right) \in \mathbb{R}_+^N$, $\forall q \in \mathcal{Q} \triangleq \{1,\ldots,Q\}$, and the overall power allocation profile for all SUs is denoted by $\boldsymbol{p} = \left(\underline{p}_1, \underline{p}_2, \ldots, \underline{p}_Q\right) \in \mathbb{R}_+^{N\times Q}$.

The received signal at SR_q in band n can be written as:

$$y_{qn} = \sqrt{p_{qn} h_{qn}} v_{qn} + \sum_{k\in\mathcal{K}} i_{qn}^{(k)} + w_{qn}, \tag{1}$$

where h_{qn} is the power gain of the direct link $ST_q - SR_q$; $v_{qn} \backsim \mathcal{CN}(0,1)$ is the normalized transmitted signal by SU q (a zero-mean circularly symmetric complex Gaussian variable of unit variance), $w_{qn} \backsim \mathcal{CN}\left(0, \sigma_{qn}^2\right)$ is the noise in band n for SU_q of variance σ_{qn}^2; and $i_{qn}^{(k)} \backsim \mathcal{CN}\left(0, \left(\tau_{qn}^{(k)}\right)^2\right)$ is the interfering signal from PU $k \in \mathcal{K} \triangleq \{1,\ldots,K\}$ of variance $\left(\tau_{qn}^{(k)}\right)^2$. Since the transmit power of the PUs cannot be impacted by the secondary system, the terms $i_{qn}^{(k)}$ are just some fixed parameters in our model. Only the knowledge of the SINRs are needed and measured at each SU. All links are assumed to be stationary and independent of the noise.

The Gaussian input, interference and noise assumptions are fairly standard in the signal processing for communications literature [3, 27, 28]. The main scientific reasons behind this are that the Gaussian noise is known to be the worst additive noise distribution in terms of the Shannon achievable rate [29], and that the Gaussian input is optimal in a Gaussian environment [30].

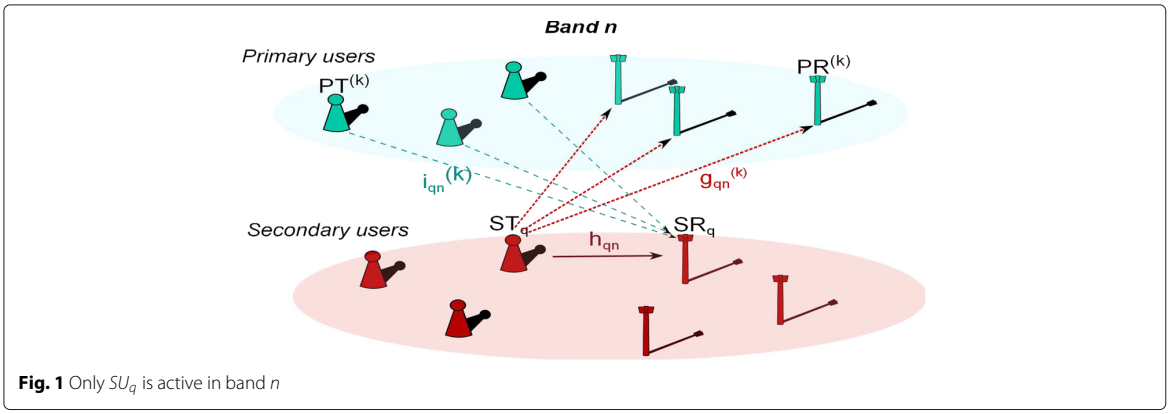


Fig. 1 Only SU_q is active in band n

In this context, we define the QoS measure for SU q in band n by the Shannon capacity expression in [31], assuming that the corresponding bandwidth is unitary:

$$c_{qn}(x_{qn}, p_{qn}) = x_{qn} \log_2 \left(1 + s_{qn} p_{qn}\right) \tag{2}$$

where s_{qn} is the signal-to-interference-plus-noise-ratio (SINR) of the direct link of q^{th} SU ($ST_q - SR_q$), $s_{qn} = h_{qn} / \left(\sigma_{qn}^2 + \sum_{k=1}^{K} \left(\tau_{qn}^{(k)}\right)^2\right)$. The spectrum allocation policy for the q^{th} SU is denoted by $\underline{x}_q = (x_{q1}, x_{q2}, \ldots, x_{qN})$, $\forall q \in \mathcal{Q}$, and the overall allocation profile for all SUs is denoted by $\boldsymbol{x} = (\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_Q)$.

This work is solely focused on the case in which at most one SU is allocated in each band n to avoid interference impairments (to the PUs and among SUs). This means that only discrete spectrum allocation policies $x_{qn} \in \{0, 1\}$ are allowed by the spectrum manager. Assuming orthogonal and unit bandwidth channels, the overall achievable rate of the q^{th} SU transmission is

$$C_q\left(\underline{x}_q, \underline{p}_q\right) = \sum_{n \in \mathcal{N}} c_{qn}(x_{qn}, p_{qn}). \tag{3}$$

The assumption that only one SU is allowed in each band is also made for tractability reasons. Indeed, optimizing the system's achievable sum-rate or some other objective under SU power or rate constraints is a very difficult, intractable and non-convex problem in an interference channel model. The achievable rate of one SU C_q is a non-convex function with respect to $\underline{p}$, because of the interference terms from the other SUs. To overcome this issue, one possibility is to decentralize the decision process and to consider a distributed cognitive radio system, in which each SU chooses its own power allocation policy to optimize its own objective [3, 6, 11]. The underlying non-cooperative game is a convex game as each SU's rate is convex w.r.t. its own controlled variables $\underline{p}_{-q}$.

However, the major disadvantage of this approach is that the resulting Nash equilibrium solution provides an operating point that performs very poorly compared to a centralized solution. Since we are interested in a centralized cognitive radio system, a second way to overcome this major difficulty (the non convexity of C_q) is to limit the access of the SUs to the spectrum; only one SU is allowed per band.

Although the presence of PUs impacts the rates of the SUs via the SINR terms s_{qn}, their crucial impact is in the additional interference constraints imposed by their presence on the SUs' transmit powers. Indeed, in our CR system, the SUs are allowed to transmit only if the created interference to the primary receivers is guaranteed to be kept below some predefined thresholds to protect the transmissions of the PUs.

3 Joint scheduling and power allocation problem

The main objective of the paper is to study a centralized resource allocation problem in which the spectrum manager wishes to schedule the SUs in an effort to minimize the overall power consumption (in coherence with the green communications sprint [32, 33]) while meeting minimum QoS constraints at the SUs and without interfering with the PUs above the critical limits. Two main questions arise:

- Spectrum scheduling: which SU should be scheduled in each band?
- Power allocation: what is the optimal power allocation policy for each SU in its allocated bands?

In order to tackle these questions, we formulate the problem as follows:

$$(DP_1) \quad \begin{aligned} &\text{minimize} \sum_{q \in \mathcal{Q}} \sum_{n \in \mathcal{N}} p_{qn} \\ &\text{s.t.} \sum_{n \in \mathcal{N}} g_{qn}^{(k)} p_{qn} \leq \overline{P}_q^{(k)}, \ \forall q, \forall k \\ &\quad 0 \leq g_{qn}^{(k)} p_{qn} \leq P_{qn}^{peak(k)}, \ \forall n, \forall q, \forall k \\ &\quad C_q\left(\underline{x}_q, \underline{p}_q\right) \geq R_q^{min}, \forall q \\ &\quad \sum_q x_{qn} \leq 1, \forall n \\ &\quad x_{qn} \in \{0, 1\}, \forall n, \forall q, \end{aligned}$$

where R_q^{min} represents the target QoS at SU_q; $g_{qn}^{(k)}$ is the power gain of the interfering link $ST_q - PR^{(k)}$; $\overline{P}_q^{(k)}$ is the maximum average interference power that SU_q is allowed to inflict on PU_k; and $P_{qn}^{peak(k)}$ is the maximum peak interference power in band n that SU_q is allowed to inflict on PU_k.

This optimization problem is difficult for two reasons. First, the target QoS constraints and the maximum interference constraints inflicted on the PUs are opposing ones and, thus, the feasible set may be void depending on the system parameters. Second, to avoid the interference impairments to the PUs and among SUs, we assume that the system owner schedules at most one SU to a given band n and that such a band cannot be further fractioned. This turns the problem into a discrete optimization with respect to the scheduling policy. In the remainder of this section, we tackle both issues and provide an efficient algorithm to compute the optimal solution when it exists.

3.1 Feasible set

The spectrum manager has to schedule all SUs to ensure a non-zero QoS target for each SU. These constraints might require the SUs to transmit at power levels which inflict an interference level that is unacceptable by the primary

system. Thus, the first arising question is under what conditions on the system parameters are the QoS and the interference constraints met simultaneously?

We denote the feasible set of (DP_1) by $\mathcal{S}_{\mathcal{F}}$:

$$\mathcal{S}_{\mathcal{F}} = \left\{ \begin{array}{l} (\boldsymbol{x},\boldsymbol{p}) \in \{0,1\}_{+}^{N\times Q} \times \mathbb{R}_{+}^{N\times Q} : \\ \sum\limits_{n\in\mathcal{N}} g_{qn}^{(k)} p_{qn} \leq \overline{P}_{q}^{(k)},\ \forall q, \forall k \\ 0 \leq g_{qn}^{(k)} p_{qn} \leq P_{qn}^{peak(k)},\ \forall n, \forall q, \forall k \\ C_q\left(\underline{x}_q, \underline{p}_q\right) \geq R_q^{min}, \forall q \\ \sum\limits_{q\in\mathcal{Q}} x_{qn} \leq 1, \forall n \text{ and } x_{qn} \in \{0,1\}, \forall n, \forall q. \end{array} \right\} \quad (4)$$

We provide next sufficient and necessary conditions on the system parameters for the existence of at least one solution to (DP_1), i.e., $\mathcal{S}_{\mathcal{F}} \neq \emptyset$.

We start by assuming, without loss of generality, that $R_q^{min} > 0$, for all SU q. Given this assumption and the fact that only one SU is allowed to transmit in a given band, the problem has no solution when $N < Q$. Thus, a first trivial necessary condition for a solution to exist is that $N \geq Q$. In this case, we derive further necessary conditions and sufficient conditions for the existence of a solution.

Theorem 1. ***(Necessary conditions)*** *Assuming that $N \geq Q$ and $R_q^{min} > 0$, $\forall q$, if the minimum rate R_ℓ^{min} of an arbitrary SU ℓ is greater than its maximum achievable rate R_ℓ^{max}, then the feasible set $\mathcal{S}_{\mathcal{F}}$ is void. Here, R_ℓ^{max} represents the optimal value of the following optimization problem:*

$$\begin{array}{l} maximize \sum\limits_{n\in\mathcal{N}} \log_2\left(1 + p_{\ell n} s_{\ell n}\right) \\ s.t.\ \sum_n g_{\ell n}^{(k)} p_{\ell n} \leq \overline{P}_{\ell}^{(k)} \\ \quad g_{\ell n}^{(k)} p_{\ell n} \leq P_{\ell n}^{peak(k)},\ \forall n, \forall k \end{array} \quad (5)$$

which corresponds to the maximum achievable rate of SU ℓ if it were the only SU in the network.

This result means that, if there is at least one SU ℓ that cannot achieve its minimum requirement R_ℓ^{min} even if it were the only SU in the system, the problem (DP_1) is infeasible. In the particular case in which there is only one SU in the system, i.e., $Q = 1$, these necessary conditions are tight as they are also the sufficient conditions guaranteeing the existence of a solution. Although these conditions are not tight when $Q > 1$ (at least one band has to be allocated per SU and, thus, no SU will have access to all bands), they have the merit of being general, fair from the SUs perspective, intuitive, and having a low computational complexity. The proof of this Theorem is detailed in Appendix A1.

Theorem 2. ***(Sufficient conditions)*** *Assuming that $N \geq Q$ and $R_q^{min} > 0$, $\forall q$, if for each SU q the minimum rate R_q^{min} is lower than the following threshold: $R_q^{CS} = \log_2\left(1 + \min\limits_n \left\{ s_{qn} \min\limits_k \left\{ \frac{\overline{P}_q^{(k)}}{g_{qn}}, \frac{P_{qn}^{peak(k)}}{g_{qn}} \right\} \right\}\right)$, then the feasible set is non-void $\mathcal{S}_{\mathcal{F}} \neq \emptyset$.*

Intuitively, if each SU_q has a minimum rate R_q^{min} small enough, i.e., smaller than the rate obtained when using only its worst channel, then the problem (DP_1) is feasible.

Remark 3.1. *When $Q = 1$, $N = 1$, the sufficient conditions of Theorem 2 are identical to the necessary conditions of Theorem 1. However, (similarly to the necessary conditions above) these sufficient conditions are not tight in general. Indeed, when $Q = 1$ and $N > 1$ the necessary conditions in Theorem 1 are the tight sufficient conditions. Nevertheless, in the most general case, it seems very difficult to find better sufficient conditions that are computationally tractable and fair with respect to all SUs. To better understand this, consider first the case in which $Q \leq N < 2Q$: there are not enough channels to allocate two channels per SU. Thus, some of the SUs will only be allocated a single channel which may very well be their worst channel. In this case, finding better and fair sufficient conditions seems very unlikely. Now, if $\gamma Q \leq N < (\gamma+1)Q$ with $\gamma \geq 2$: at least $\gamma \geq 2$ channels may be allocated to each SU. In this case, a better sufficient condition could be found by computing the achievable rate over the worst γ channels for each SU. However, since we cannot know in advance which combination of γ channels results in a worst case achievable rate, one would have to compute all C_N^γ combinations for each of the Q SUs. The complexity of such an approach is therefore prohibitive.*

The proof of this Theorem is detailed in Appendix A2.

Now, if the feasible set $\mathcal{S}_{\mathcal{F}}$ is non-void, finding the solution of (DP_1) is not trivial. Indeed, we notice that the scheduling of SUs is a discrete combinatorial problem $x_{qn} \in \{0,1\}$, $\forall n$, $\forall q$. In the following, we will provide an efficient algorithm that computes the optimal solution in a very efficient manner using a Lagrangian relaxation approach [14].

3.2 Lagrangian relaxation

As we have already mentioned, we assume that at most one SU is allowed to transmit in a given band. The constraints on the discrete scheduling policy $\boldsymbol{x}$ make the problem (DP_1) very difficult to solve in this form. In the following, we will use a Lagrangian relaxation technique to overcome this issue.

When the feasible set $\mathcal{S}_{\mathcal{F}}$ is not void, we propose to solve a continuous problem in which the scheduling parameter

is continuous $x_{qn} \in [0,1]$. This relaxed problem (CP_2) holds the advantage of being a convex optimization problem, much simpler to solve than (DP_1):

$$(CP_2)\quad \begin{aligned} &\text{minimize} \sum_{q\in\mathcal{Q}}\sum_{n\in\mathcal{N}} p_{qn} \\ &\text{s.t.} \sum_{n\in\mathcal{N}} g_{qn}^{(k)} p_{qn} \le \overline{P}_q^{(k)},\ \forall q, \forall k \\ &0 \le g_{qn}^{(k)} p_{qn} \le x_{qn} P_{qn}^{peak(k)},\ \forall n, \forall q, \forall k \\ &R_q\left(\underline{x}_q, \underline{p}_q\right) \ge R_q^{min}, \forall q \\ &\sum_q x_{qn} \le 1, \forall n \\ &0 \le x_{qn} \le 1, \forall n, \forall q. \end{aligned}$$

Inspired by [13], the continuous problem (CP_2) is not exactly the Lagrangian relaxation of (DP_1) but that of an equivalent discrete problem. Two differences can be observed in (CP_2). First, the peak interference power constraints $g_{qn}^{(k)} p_{qn} \le P_{qn}^{peak(k)}$ are replaced by $g_{qn}^{(k)} p_{qn} \le x_{qn} P_{qn}^{peak(k)}$. Second, we replace the Shannon achievable rate of SU q by the function:

$$R_q(\underline{x}_q, \underline{p}_q) = \sum_{n\in\mathcal{N}} r_{qn}(x_{qn}, p_{qn}), \tag{6}$$

where

$$r_{qn}(x_{qn}, p_{qn}) = \begin{cases} x_{qn} \log_2\left(1 + \frac{s_{qn} p_{qn}}{x_{qn}}\right), & \text{if } x_{qn} > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

Remark 3.2. *Our target problem (DP_1) focuses only on discrete spectrum allocation policies $x_{qn} \in \{0,1\}$. Under this assumption, the function in Eq.* (7) *corresponds exactly to the achievable Shannon rate in* (2)*. However, in the continuous case in which $x_{qn} \in [0,1]$ expression* (7) *does not correspond to the Shannon achievable rate. The denominator of the term $\frac{s_{qn} p_{qn}}{x_{qn}}$ is introduced for technical purposes and plays an important role in proving the optimality of the Lagrangian relaxation approach when solving (DP_1).*

The major interest of this approach is that it will allow us to solve the initial non-convex problem using a very efficient algorithm based on convex optimization techniques. Not only is the continuous problem (CP_2) simpler to solve, but also, by using these two modifications, the problem (CP_2) will have discrete solutions in $\boldsymbol{x}$ which will turn out to be the solutions to our initial problem (DP_1), as we will show later on.

In the following, we will give some proprieties of the problem (CP_2) where we denote by $(\boldsymbol{x}^*, \boldsymbol{p}^*)$ the pair of scheduling and power allocation that satisfies all the constraints at the optimum:

Proposition 1. *In the continuous problem (CP_2), all the rate constraints are active at the optimal solution, i.e., the rate constraints are satisfied with equality.*

This result means that if the feasible set is non-void, the optimal rate at each SU achieves the target QoS R^{min}.

Proposition 2. *In the continuous problem (CP_2), all the scheduling average constraints are all active at the optimal solution : $\sum_q x_{qn}^* = 1$, $\forall q$.*

This result means that if the feasible set is non-void, all bands are fully used by the opportunistic users. The proofs of these two propositions are detailed in Appendix B.

3.2.1 Dual formulation

The continuous problem (CP_2) is a convex optimization problem. Indeed, the objective function $\sum_q \sum_n p_{qn}$ is affine in the overall power allocation profile $\boldsymbol{p} = \left(\underline{p}_1, \underline{p}_2, \ldots, \underline{p}_Q\right) \in \mathbb{R}_+^{N\times Q}$ and, regarding the constraints, the interference constraints are both affine in (x,p); the scheduling constraints are affine in x; and the rate constraint is jointly concave in (x,p). From the convex optimization problem definition [34], it follows that (CP_2) is convex and, thus, can be solved via a dual formulation. We associate dual variables $\boldsymbol{\lambda} = (\lambda_q)_{q\in\mathcal{Q}}$ with total interference power constraints, $\boldsymbol{\beta} = (\beta_q)_{q\in\mathcal{Q}}$ with rate constraints and $\boldsymbol{\mu} = (\mu)_{n\in\mathcal{N}}$ with scheduling constraints, resulting in the following Lagrangian:

$$\begin{aligned} \mathrm{L}(\boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\mu}, \mathbf{x}, \mathbf{p}) &= \sum_q \sum_n p_{qn} - \sum_q \beta_q \left(R_q\left(\underline{x}_q, \underline{p}_q\right) - R_q^{min}\right) \\ &+ \sum_k \sum_q \lambda_q^{(k)} \left(\sum_n g_{qn}^{(k)} p_{qn} - \overline{P}_q^{(k)}\right) + \sum_n \mu_n \left(\sum_q x_{qn} - 1\right). \end{aligned} \tag{8}$$

To solve (CP_2) it suffices to solve:

$$\max_{(\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\mu})\ge 0} \min_{(\mathbf{x},\mathbf{p})\in\mathcal{S}_{\mathcal{F}}} \mathrm{L}(\boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\mu}, \mathbf{x}, \mathbf{p}). \tag{9}$$

Inspired by [13], in which the authors consider a rate maximization joint scheduling and power allocation problem in OFDM systems without the PU constraints, we propose to solve here a more complex problem (CP_2) via the following steps.

Step 1: The optimal power $\mathbf{p}$ which minimizes $\mathrm{L}(\boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\mu}, \mathbf{x}, \mathbf{p})$ given fixed $\boldsymbol{\lambda}$, $\boldsymbol{\beta}$ and $\boldsymbol{\mu}$[1] is a water-filling type of solution:

$$p_{qn}^* = \frac{x_{qn}}{s_{qn}} \left[\frac{\beta_q s_{qn}}{1 + \sum_k \lambda_q^{(k)} g_{qn}^{(k)}} - 1\right]_0^{\min_k \left\{P_{qn}^{peak(k)}\right\} s_{qn}} \tag{10}$$

Step 2: Substituting $\mathbf{p}^*$ into $\mathrm{L}(\boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\mu}, \mathbf{x}, \mathbf{p})$ yields to the following affine function in $\mathbf{x}$:

$$\begin{aligned}\mathrm{G}(\boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\mu}, \mathbf{x}) = &-\sum_q \sum_n x_{qn}\left(\varphi_{qn} - \mu_n\right) \\ &+\sum_q \beta_q R_q^{min} - \sum_k \sum_q \lambda_q^{(k)} \overline{P}_q^{(k)} - \sum_n \mu_n,\end{aligned} \tag{11}$$

where $\varphi_{qn} = \varphi_{qn}(a, b, c, d)$ is a function of the system parameters defined by:

$$\begin{cases} 0 & \text{if } \frac{a}{c} \geq b \\ -b + \frac{a}{c} + b\log_2\left(\frac{bc}{a}\right) & \text{if } \frac{a}{c} \leq b \leq \frac{a(dc+1)}{c} \\ -da + b\log_2(1+dc) & \text{if } \frac{(a)(dc+1)}{c} < b \end{cases} \tag{12}$$

for $a = 1 + \sum_k \lambda_q^{(k)} g_{qn}^{(k)}$, $b = \beta_q$, $c = s_{qn}$ and $d = \min_k \left\{ P_{qn}^{peak(k)} \right\}$.

Step 3: From Eq. (11), we remark that this function is affine in $\boldsymbol{x}$. Optimizing it over $\boldsymbol{x}$ such that $x_{qn} \in [0, 1]$ yields the dual function:

$$\begin{aligned}\mathrm{G}(\boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\mu}) = &-\sum_q \sum_n \left[\varphi_{qn} - \mu_n\right]^+ + \sum_q \beta_q R_q^{min} \\ &- \sum_k \sum_q \lambda_q^{(k)} \overline{P}_q^{(k)} - \sum_n \mu_n,\end{aligned} \tag{13}$$

where the optimal scheduling allocation $\boldsymbol{x}^*$ is:

$$x_{qn}^*(\boldsymbol{\mu}) = \begin{cases} 1 \text{ if } \varphi_{qn} > \mu_n \\ 0 \text{ if } \varphi_{qn} < \mu_n \end{cases} \tag{14}$$

and if $\varphi_{qn} = \mu_n$, then the optimal value x_{qn}^* can be anything in the interval $[0, 1]$. In such cases and from Proposition 2, one must only ensure that all scheduling constraints are met: $\sum_q x_{qn}^* = 1$ for all n. We can now maximize the dual function $\mathrm{G}(\boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\mu})$ over $\boldsymbol{\mu}$ for given $\boldsymbol{\lambda}$ and $\boldsymbol{\beta}$, by setting $\mu_n = \mu_n^*(\boldsymbol{\lambda}, \boldsymbol{\beta})$ similarly to [13] where μ_n^* is obtained as follows:

$$\mu_n^*(\boldsymbol{\lambda}, \boldsymbol{\beta}) = \max_q \varphi_{qn}(\lambda_q^{(k)}, \beta_q). \tag{15}$$

Remark 3.3. *From Eqs.* (14) *and* (15)*, it is clear that* $x_{qn}^*(\boldsymbol{\mu}) = 0$ *if* $q \notin arg\max_{q \in \mathcal{Q}} \varphi_{qn}\left(\lambda_q^{(k)}, \beta_q\right)$*. Intuitively, this means that band n is allocated to a specific SU* ℓ *if it maximizes a specific channel metric given by* $\varphi_{\ell n}\left(\lambda_\ell^{(k)}, \beta_\ell\right)$*. There may be ties when multiple SUs achieve the value* μ_n^* *in band n. For example, if there exist two SUs r, q such that* $r \neq q$ *and* $\varphi_{qn} = \varphi_{rn} = \mu_n^*$*. These ties happen with zero probability if the independent random channel gains are drawn from continuous distributions. In practice, this implies that such problematic cases almost never occur.*

From Eq. (14) we also observe that (CP_2) allows for discrete solutions in $\boldsymbol{x}^*$ when it is feasible. Since we are interested in solving the discrete problem (DP_1), we make the choice to select a discrete solution such that

$$x_{qn}^*(\boldsymbol{\mu}^*) = \begin{cases} 1 \text{ if } \varphi_{qn} = \mu_n^* \\ 0 \text{ if } \varphi_{qn} < \mu_n^*. \end{cases} \tag{16}$$

Step 4: Substituting $\boldsymbol{\mu}^*$ into $\mathrm{G}(\boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\mu})$ and noticing that $\boldsymbol{\mu}^*, \mathbf{p}^*, \mathbf{x}^*$ are all functions of $(\boldsymbol{\lambda}, \boldsymbol{\beta})$ we further have:

$$\begin{aligned}\mathrm{G}(\boldsymbol{\lambda}, \boldsymbol{\beta}) = &-\sum_n \max_q \varphi_{qn}\left(\lambda_q^{(k)}, \beta_q\right) + \sum_q \beta_q R_q^{min} \\ &- \sum_k \sum_q \lambda_q^{(k)} \overline{P}_q^{(k)}.\end{aligned} \tag{17}$$

The solution to the dual problem (9) can be computed numerically by maximizing $\mathrm{G}(\boldsymbol{\lambda}, \boldsymbol{\beta})$ over $\boldsymbol{\lambda} \geq 0$ and $\boldsymbol{\beta} \geq 0$. To this aim, we use a sub-gradient based search [14] to update $\boldsymbol{\lambda}$ and $\boldsymbol{\beta}$.

3.2.2 Projected sub-gradient algorithm

Based on the previous dual formulation, we propose an iterative sub-gradient algorithm to compute the solution of (CP_2) when it exists. The iterations are detailed in Algorithm 1. The sub-gradient approach is usually employed to compute water-filling type of solutions [3, 10, 13]. The proposed algorithm converges to the optimal solution whenever it is feasible.

Other efficient algorithms such as interior point methods, which are based on Newton's iteration and on second-order derivatives, can be implemented to solve (CP_2). Both types of algorithms have advantages and inconveniences, and their convergence performance depends on the compromise between the number of iterations and the complexity of each iteration. The sub-gradient method has the advantage of being simple as its iterations are of low-complexity and scale as $O(Q \times (K + N))$.

Proposition 3. *If* (CP_2) *is feasible, then Algorithm 1 always converges and the convergence point is an optimal scheduling and power allocation policy.*

The proof of this proposition is based on convex optimization tools and is given in Appendix C. The intuition is that, at the convergence state, the iteration, which is discrete with respect to the spectrum allocation policy by construction, satisfies the KKT conditions which are both necessary and sufficient for optimality in convex optimization problems.

3.3 Optimality of the Lagrange relaxation

Our main objective in the remainder of the paper is to show that Algorithm 1 not only converges to the solution of (CP_2) but to that of the initial discrete problem (DP_1).

Let us define by DP_2 the discrete version of (CP_2). The objective of DP_2 is to minimize the SU overall power consumption subject to the rate constraints, the interference power constraints (identical to the ones in (CP_2)) and the discrete scheduling constraints: $x_{qn} \in \{0,1\}$, $\forall n, \forall q$, $\sum_q x_{qn} \leq 1$, $\forall n$. Because of space limitations, we do not write the expressions of DP_2 explicitly.

Proposition 4. *If (CP_2) is feasible, then all the discrete solutions of (CP_2) (with respect to the optimal scheduling allocation $\boldsymbol{x}^*$) are the only optimal solutions of DP_2.*

Algorithm 1: Projected Sub-gradient Algorithm for the Joint Scheduling and Power Allocation Problem

1) Initialization: $\beta_q^{[0]}$, $\lambda_q^{(k)[0]}$, $\forall\, q$, $\forall\, k$
2) $t = 1$
3) while *iteration* t **do**
  for *each SU* $q \in \mathcal{Q}$ **do**
    for *each Band* $n \in \mathcal{N}$ **do**
      Update powers $p_{qn}^{[t]}$ given in (10)
      Update channel metrics φ_{qn} defined in (12)
      Update the Lagrange multiplier $\mu_n^{[t]}$ using (15)
      Update the scheduling allocation $x_{qn}^{[t]}$ using (16)
    end
    for $k \in \mathcal{K}$ **do**
      Update $\lambda_q^{(k)[t]}$ such that[2]:
      $$\lambda_q^{(k)[t]} = \left[\lambda_q^{(k)[t-1]} - \tau\left(\overline{P}_q^{(k)} - \sum_n g_{qn}^{(k)} p_{qn}^{[t]}\right)\right]^+$$
    end
    Update $\beta_q^{[t]}$ such that:
    $$\beta_q^{[t]} = \left[\beta_q^{[t-1]} + \tau\left(R_q^{min} - R_q\left(\underline{x}_q^{[t]}, \underline{p}_q^{[t]}\right)\right)\right]^+$$
    where R_q is defined in (6)
    $t = t + 1$
  end
end
4) Repeat **3)** until convergence for all SUs in all the bands $\forall q, \forall n$, $\left|x_{qn}^{[t]} - x_{qn}^{[t-1]}\right| \leq \epsilon_x$, $\left|p_{qn}^{[t]} - p_{qn}^{[t-1]}\right| \leq \epsilon_p$, $\left|\beta_q^{[t]} - \beta_q^{[t-1]}\right| \leq \epsilon_\beta$ and
$\max_{k\in\mathcal{K}} \left\{\left|\lambda_q^{(k)[t]} - \lambda_q^{(k)[t-1]}\right| - \epsilon_{\lambda_k}\right\} \leq 0$
where ϵ_x, ϵ_p, ϵ_{λ_k} and ϵ_β are precision parameters.

Proof. The feasible set of DP_2 is included in the feasible set of (CP_2). The only difference between both problems is the scheduling parameter x_{qn} which is discrete in DP_2 and continuous in (CP_2). On the one hand, we can see that if there are any discrete solutions to (CP_2), then these discrete solutions will also be the optimal solutions of DP_2. On the other hand, if the solutions of (CP_2) are continuous (i.e., the scheduling allocations are between 0 and 1), then the SUs' power allocations policies are constants (as they do not depend on the scheduling allocations), thus, these continuous solutions have the same efficiency comparing to the discrete solutions on the borders (i.e., are equal to 0 or 1). □

Next, we compare the discrete problems (DP_1) and DP_2.

Proposition 5. *The optimization problems (DP_1) and DP_2 are equivalent in the sense that their solution sets are identical.*

Proof. If we fix an arbitrary scheduling policy $\widetilde{\boldsymbol{x}} \in \mathcal{X}$, where the scheduling constraints set is defined as

$$\mathcal{X} = \left\{\widetilde{\boldsymbol{x}} \in \{0,1\}^{Q\times N} \mid \sum_l \widetilde{x}_{ln} = 1, \forall n\right\}$$

then, both problems (DP_1) and DP_2 reduce to the same power allocation problem (CP_3) below in which the only variable left is the power allocation policy[3] at the SU's

$$(CP_3)\quad \begin{aligned} &\text{minimize} \sum_{q\in\mathcal{Q}}\sum_{n\in\mathcal{N}} p_{qn}\mathbb{1}_{\{\widetilde{x}_{qn}\}} \\ &\text{s.t.} \sum_n g_{qn}p_{qn}\mathbb{1}_{\{\widetilde{x}_{qn}\}} \leq \overline{P}_q,\ \forall q \in \mathcal{Q} \\ &\quad 0 \leq g_{qn}p_{qn}\mathbb{1}_{\{\widetilde{x}_{qn}\}} \leq P_{qn}^{peak}\mathbb{1}_{\{\widetilde{x}_{qn}\}}, \forall q, n \\ &\quad \sum_q \mathbb{1}_{\{\widetilde{x}_{qn}\}} \log_2\left(1 + \frac{s_{qn}p_{qn}}{\widetilde{x}_{qn}}\right) \geq R_q^{min},\ \forall n. \end{aligned}$$

On the one hand, if the scheduling parameter equals zero, $\widetilde{x}_{qn} = 0$, then the corresponding optimal power is also zero $\widetilde{p}_{qn} = 0$ (because the optimal power is linear in $\boldsymbol{x}$). On the other hand, if the scheduling parameter equals one, $\widetilde{x}_{qn} = 1$, then the optimal powers $\widetilde{p}_{qn}$ are given by a water-filling type of solution over the allocated spectrum. Thus, whenever $\widetilde{\boldsymbol{x}}$ is fixed, both problems (DP_1) and DP_2 reduce to solving the same Q decoupled power allocation problems.

It remains to prove that both DP_2 and (DP_1) have the same set of optimal spectrum allocation policies $\boldsymbol{x}^*$. The optimal powers $\boldsymbol{p}^*$ will follow by solving (CP_3). Assuming that (DP_1) is feasible, we can easily check that the optimal

solutions of (DP_1) meet all the constraints of DP_2 as well. Adding the fact that the feasible set of (DP_1) contains (and is larger) than the feasible set of DP_2, we conclude that the optimal solution sets of the two problems are identical. □

Theorem 3. *If (CP_2) is feasible, then all solution of (CP_2) that are discrete w.r.t $\boldsymbol{x}^*$ are the optimal solutions of (DP_1).*

This Theorem follows from Propositions 4 and 5. In Proposition 5, we have shown that both discrete problems DP_2 and (DP_1) are equivalent. Moreover, from Proposition 4, we have shown that the optimal discrete solutions of DP_2 are the discrete solutions of (CP_2). Thus, we can conclude that all discrete solutions of (CP_2) are the optimal discrete solutions of (DP_1).

Corollary 1. *The feasibility of the relaxed problem (CP_2) implies the feasibility of (DP_1).*

This is important because the feasibility of (CP_2) problem, which is a convex problem, is much simpler to study than the feasibility of the discrete problem (DP_1). In other words, to decide whether the initial problem (DP_1) is convex, we have simply to solve the problem of feasibility of (CP_2) which is convex.

In conclusion, our results show that Algorithm 1, initially built to solve (CP_2), selects only the optimal discrete solutions and actually solves DP_2. From Proposition 3, both discrete problems DP_2 and (DP_1) are equivalent; thus, our projected sub-gradient Algorithm 1 solves the initial problem (DP_1) and converges to the optimal scheduling and power allocation policy whenever the problem is feasible.

In the following section, we will present some simulation results which illustrate the performance of our projected sub-gradient algorithm.

4 Numerical results

All observations in this section have been verified via extensive simulations with generic system parameters. We have selected only a few of the most illustrative and interesting results to be presented next.

4.1 Power consumption efficiency

We start by comparing the overall power consumption between our optimal scheduling policy and arbitrarily scheduling techniques such as interleaved and block-wise spectrum scheduling [35]. Once the spectrum scheduling is fixed, computing the power allocation policies follows via water-filling type of sub-gradient methods. We focus only on the cases in which the problem is feasible and in which our Algorithm 1 converges to the optimal solution of (DP_1). Thus, for each of 10^4 random channel realizations, we choose the minimum target QoS that are equal to the rates in the sufficient conditions of Theorem 2.

Figure 2 illustrates this comparison as function of $N \in \{Q, 2Q, 4Q, 8Q, 16Q\}$ in the scenario: [4] $Q = 4$, $K = 8$, the channel gains are drawn randomly $g_{qn}^{(k)} \backsim \mathcal{N}(0,4)$, $\forall q, \forall n, \forall k$ and $s_{qn} \backsim \mathcal{N}(0,20)$, $\forall q, \forall n$, $\overline{P}_q^{(k)} = 10$ mW, $\forall q, \forall k$, $P_{qn}^{peak(k)} = 20$ mW, $\forall q, \forall n, \forall k$.

We remark that the interleaved and block-wise allocations have the same average performance because of the independent and identically distributed (i.i.d) channel gains. Our algorithm outperforms these two fixed-spectrum allocations (interleaved or block-wise allocation), and the performance gap decreases with N.

The sufficient conditions guaranteeing the existence of a solution and the convergence of Algorithm 1 are not tight in general as they rely on the use of Q bands alone and on the assumption that each SU is allowed to transmit in its worst channel. Finding tighter sufficient conditions that are tractable seems a very difficult task. Therefore, we study the performance of our algorithm in the cases in which the sufficient conditions are not met.

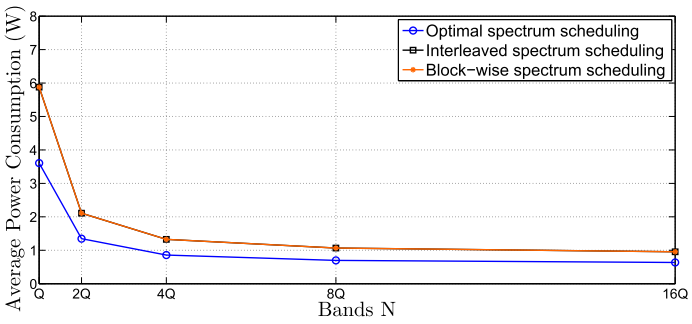


Fig. 2 Average power consumption vs. number of bands $N \in \{Q, 2Q, 4Q, 8Q, 16Q\}$ for $Q = 4$ when the sufficient conditions are always met. Our optimal spectrum scheduling outperforms the two others spectrum scheduling allocations (interleaved and block-wise)

4.2 Problem feasibility

Figure 3 illustrates the empirical probability (over 10^4 random channels realizations) that the problem (DP_1) is not feasible as function of the target QoS R^{min} in the following scenario: $N = 4$, $K = 8$, $Q = 2$, $R_q^{min} = R^{min}$, $\forall q$, $\overline{P}_q^{(k)} = 10$ mW, $\forall q, \forall k$, $\boldsymbol{s} = \begin{pmatrix} 27.8797 & 2.0727 & 0.7779 & 4.3263 \\ 7.6688 & 7.6722 & 11.0049 & 1.7281 \end{pmatrix}^4$, $P_{qn}^{peak(k)} = 20$ mW, $\forall q, \forall n, \forall k$ and the interfering channel gains are drawn randomly such that $g_{qn}^{(k)} \backsim \mathcal{N}\left(0, \sigma_g^2\right)$, $\forall q, \forall n, \forall k$ for an arbitrary chosen value $\sigma_g^2 = 2$.

In order to decide whether the problem is feasible or not, we test if the algorithm converges before the maximum number of iterations is reached. From Proposition 3 and Theorem 3, if the algorithm converges, then the convergence point is an optimal solution of (DP_1). For practical reasons, if the algorithm reaches the maximum number of iterations before convergence, we decide that the problem is not feasible. This approach is based on an empirical search for the maximum number of iterations. In this setting, we fix 10^5 maximum number of iterations.

In Fig. 3, we remark that there is a threshold, $R^{min} \simeq 6.25$ bps, below which a solution exists and above which the problem is not feasible. The other two plotted thresholds, CS and CN, illustrate the worst case sufficient conditions and necessary conditions over all random realizations and are computed as follows. We denote by $R_q^{CS(t)}$ the sufficient condition rates of Theorem 2 for the random channel realization $t \in \{1, \ldots, 10^4\}$. Similarly, we denote the necessary condition rates by $R_q^{CN(t)}$ of Theorem 1 which also depend on the random realization t. The threshold CS represents the minimum value of the rates $R_q^{CS(t)}$ over all SUs and over all random channel realizations: $CS = \min_{q,t} \left\{ R_q^{CS(t)} \right\} = 1.44$ bps and the threshold CN represents the maximum value of the rates $R_q^{max(t)}$ over SUs and over the random channel realizations: $\max_{q,t} \left\{ R_q^{CN(t)} \right\} = 12.86$ bps.

Although the CS and CN values in Fig. 3 are worst case conditions (over all random channels and all SUs), they still show that our sufficient and necessary conditions are not tight in general, as discussed in Section 3.1. This means that there are a lot of cases in which we do not know a priori whether the problem is feasible or not: all cases in which sufficient conditions are not met but the necessary conditions are met.

4.3 Necessary conditions are not met

We illustrate now the case in which the necessary conditions are not met, i.e., when there exists at least one SU q such that the target rate is above the maximum achievable rate in Theorem 1 : $R_q^{min} \geq R_q^{max}$. We consider the following setting [4]: $N = 4$, $K = 8$, $Q = 2$, $\overline{P}_q^{(k)} = 10$ mW, $\forall q, \forall k$, $\boldsymbol{s} = \begin{pmatrix} 27.8797 & 2.0727 & 0.7779 & 4.3263 \\ 7.6688 & 7.6722 & 11.0049 & 1.7281 \end{pmatrix}$, $\boldsymbol{g}^{(k)} = \begin{pmatrix} 2.4086 & 3.2329 & 1.1983 & 0.4016 \\ 1.8904 & 0.1510 & 0.0362 & 0.3318 \end{pmatrix}^5$. We fix the target rates $R_1^{min} = R_2^{min} = 13$ bps which are higher than the maximum rates that each SU could achieve if they were alone in the system: $R_1^{max} = 8.0803$ bps, $R_2^{max} = 12.8686$ bps.

Figure 4 illustrates the evolution of the Lagrange multipliers μ_n and β_q corresponding to the average constraints over the algorithm's iterations. We remark, that Algorithm 1 does not converge before the maximum number of iterations is reached. In this case, we decide that the problem is not feasible. An alternative would be to further increase the maximum number of iterations, but this process has to be finite and a pragmatic decision has to be made at some point.

4.4 Sufficient conditions are not met although necessary conditions are met

Here, we illustrate the case in which the necessary conditions are met and at least one of the sufficient conditions is not met in the same setting as Fig. 4 except for the target

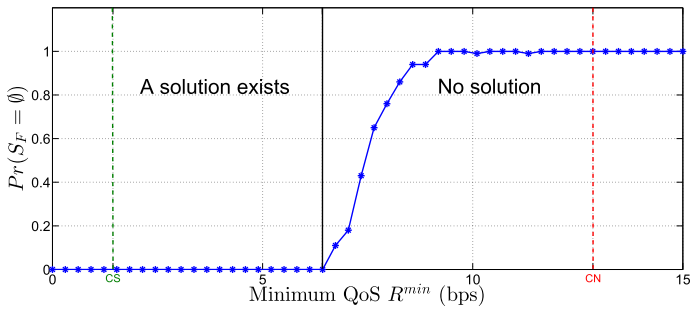


Fig. 3 Empirical probability that the problem is not feasible as function of the minimum target rate R^{min} for each SU

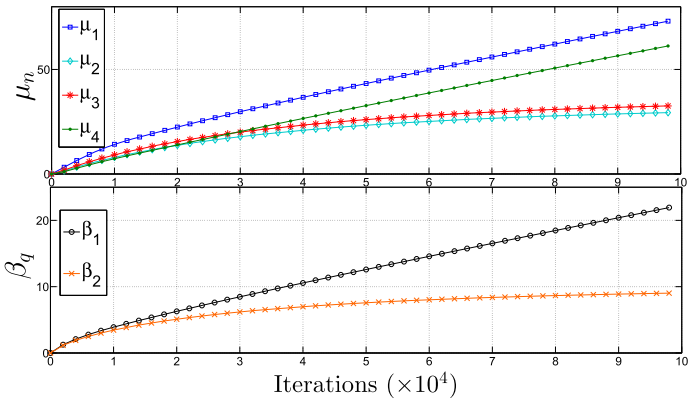


Fig. 4 Algorithm 1 reaches the maximum number of iterations (10^5) before convergence

rates which are in between the thresholds of the sufficient conditions and the necessary conditions.

4.4.1 Convergence of Algorithm 1

In Fig. 5, we want to check whether our projected subgradient algorithm converges to an optimal solution or not.

In Fig. 5a, we plot the Lagrange multipliers μ_n, β_q of the scheduling and target rate constraints when the minimum target rates are $R_1^{min} = R_2^{min} = 3$ bps. We can see that our algorithm converges in this case and, thus, the problem is feasible and solved via Algorithm 1. Indeed, the Lagrange multipliers converge to the optimal values μ_n^* and β_q^* that are strictly positive, and the scheduling and rate constraint are satisfied with equality.

In Fig. 5b, we fix the minimum target rates $R_1^{min} = R_2^{min} = 8$ bps. In this case, we remark that our algorithm reaches the maximum number of iterations. Because of such high target rates, our algorithm does not converge and, thus, the problem has no solution.

4.5 Water-filling solution

Figure 6 illustrates the optimal scheduling and power allocation policies in the case in which our algorithm converges for the same scenario as in Fig. 5a.

We remark that the optimal solution is fair in terms of bands per SU; two bands are allocated to each SU. Band 1 and 4 are allocated to SU_1, since SU_1 has higher SNR than SU_2 in these bands. Band 2 and 3 are allocated to SU_2 for the same reason. This power and spectrum allocation policy satisfies all KKT conditions, thus, it is the optimal solution.

In conclusion, when the sufficient conditions are not met although the necessary conditions are met, we do not know a priori if Algorithm 1 converges or not. Nevertheless, we can still exploit our algorithm to decide whether the problem is feasible or not and to compute the optimal solution when it exists. If the algorithm converges before reaching the maximum number of iterations, then we know that the problem is feasible and that the convergence point is an optimal solution to (DP_1). Otherwise, we can increase the maximum number of iterations or decide that the problem is not feasible and has no solution. In such cases, instead of not allowing any SU to transmit, the spectrum manager (instead of not allowing any opportunistic user to transmit) can decrease the minimum SU target rates (lower QoS may be better than no QoS in low data rate applications) or may even decide to schedule only a subset of the SUs and turn the problem into a feasible one [8].

As we have seen before, there exist other cases in which the Algorithm 1 does not calculate the solution. This happens in the indecision cases between several SUs (i.e, to decide whether the SU is allowed to transmit in a given band or not). In the following, we will try to solve some of these specific cases.

5 Particular case: symmetric systems

In this section, we are interested in a specific symmetric system in which Algorithm 1 is not suitable to compute the solution. This happens whenever ties arise between several SUs when deciding on the SU to be allocated a given band: if several SUs have the same maximum value for the decision parameter μ_n^* described in Section 3.2.1 (there exist two SUs $q \neq r$ such that $\varphi_{qn} = \varphi_{rn} = \mu_n^*$).

Consider the completely symmetric system in which all SUs experience the same channel gains, the same peak and total interference constraints, and have the same target rates: $s_{qn} = s$, $g_{qn} = g$, $P_{qn}^{peak} = P^{peak}$, $\overline{P}_q = \overline{P}$ and $R_q^{min} = R^{min}$, $\forall n$, $\forall q$.

Remark 5.1. *In this case, the optimal power allocation is such that every SU uniformly allocates its power over its allocated bands. For any scheduling policy* $\boldsymbol{x}$*, the optimal*

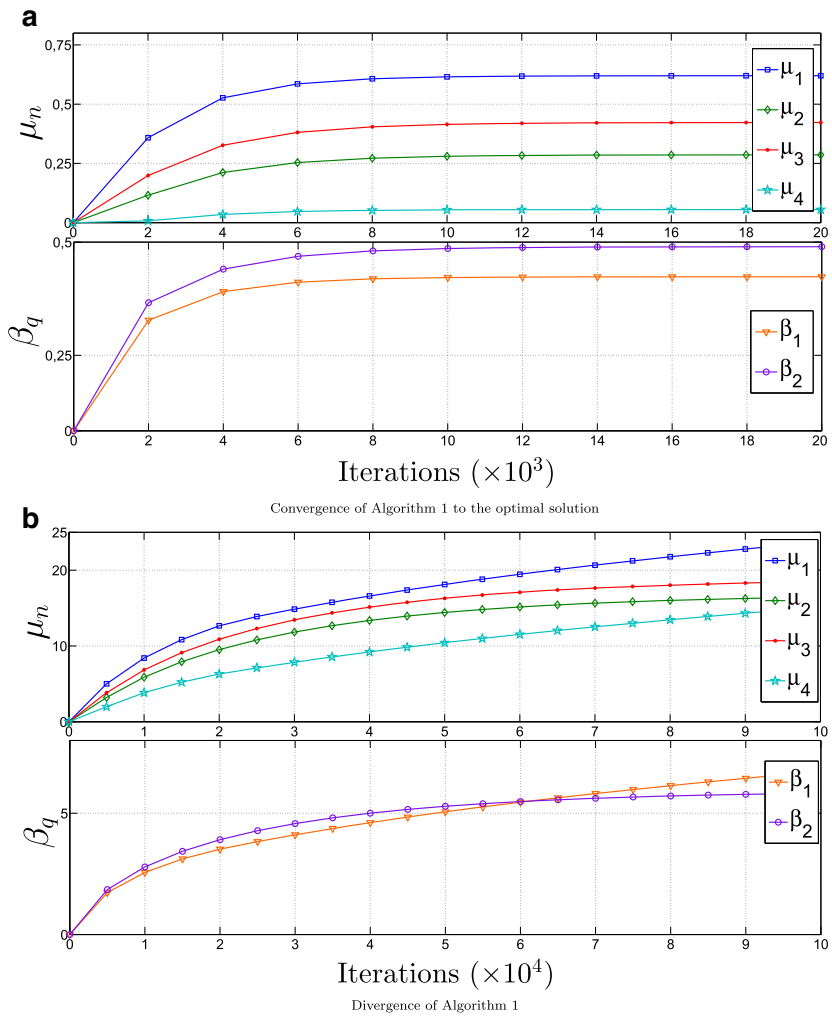


Fig. 5 Behaviour of Algorithm 1 when sufficient conditions are not met in Scenario 2: (**a**) $R^{min} = [3\ 3]$ the problem is feasible and the algorithm converges and (**b**) $R^{min} = [8\ 8]$ the algorithm does not converge in this case

power allocation is a water-filling type solution [11, 12] derived from the KKT conditions:

$$p_{qn}^{*} = \begin{cases} \frac{1}{s}\left[\frac{\beta_q s}{1+\lambda_q g} - 1\right]_0^{p^{peak}s} & if\, x_{qn}^{*} = 1 \\ 0 & otherwise \end{cases} \tag{18}$$

and, thus, we have $p_{qn}^{*} = p_q^{*}$, for any band n that is used at the optimal solution.

In the following, the main idea is to simplify the problem (DP_1) in which the unknowns are $\boldsymbol{x}$ and $\boldsymbol{p}$ and reduce it into a problem in which the only unknown is $N_q^{*}, \forall q$ where N_q^{*} denotes the number of bands allocated at the optimum to the SU_q[6] :

$$N_q^{*} \triangleq Card\left\{n \in \mathcal{N} \mid x_{qn}^{*} = 1\right\}.$$

From Proposition 2, since the scheduling constraint is satisfied with equality, the overall number of bands allocated to the SUs, at the solution, equals the total number of bands N.

$$\sum_q N_q^{*} = N. \tag{19}$$

According to Proposition 1 (rate constraint is met with equality) and Remark 5.1 (uniform power allocation is optimal), we can write the optimal power as:

$$p_{qn}^{*} = \begin{cases} \frac{1}{s}\left(\exp\left(\frac{R^{min}}{N_q}\right) - 1\right) & \text{if } x_{qn}^{*} = 1 \\ 0 & \text{otherwise.} \end{cases}$$

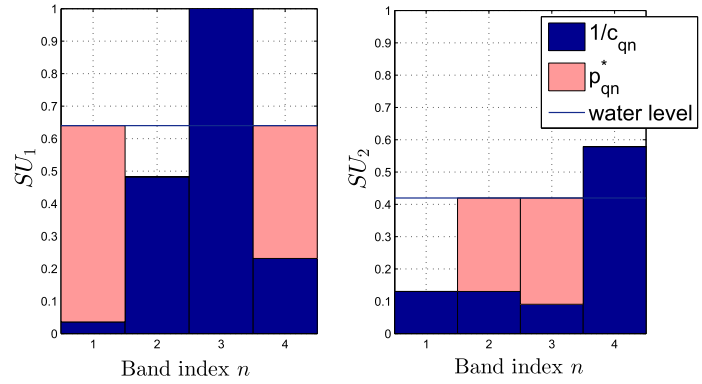


Fig. 6 Optimal scheduling and power allocation when Algorithm 1 converges. The system owner allocates two bands per SU

Using (19), the discrete problem (DP_1) simplifies as follows:

$$
(DP_N)\quad
\begin{aligned}
&\text{minimize} \sum_{q\in\mathcal{Q}} \frac{N_q}{s}\left(\exp\left(\frac{R^{min}}{N_q}\right)-1\right)\\
&\text{s.t.}\ \frac{N_q}{s}\left(\exp\left(\frac{R^{min}}{N_q}\right)-1\right) \le \frac{\overline{P}}{g},\ \forall q \quad (C1)\\
&\quad \frac{1}{s}\left(\exp\left(\frac{R^{min}}{N_q}\right)-1\right) \le \frac{P^{peak}}{g},\ \forall q \quad (C2)\\
&\quad \sum_{q=1}^{Q} N_q = N \quad (C3)\\
&\quad N_q \in \mathbb{N}^*,\ \forall q. \quad (C4)
\end{aligned}
$$

This problem is interesting as it is no longer a joint scheduling and power allocation problem $(\boldsymbol{x},\boldsymbol{p})$ and depends only on the number of allocated bands at the optimum $N_q^*,\ \forall q$.

For simplicity reasons, we start by assuming that the number of available bands is proportional to the number of SUs. This means that the ratio N/Q is an integer: $N/Q \in \mathbb{N}$.

Now, we will prove that the optimal solution to problem (DP_N) is to allocate the spectrum in a fair way to the SUs.

Proposition 6. *The optimal solution of the problem (DP_N), when feasible, is to uniformly allocate the spectrum to the SUs:*

$$N_q^* = \frac{N}{Q},\ \forall q. \tag{20}$$

The proof of this Proposition is detailed in Appendix D. The closed-form solution of the power allocation policy is given by:

$$p_{qn}^* = \begin{cases} \frac{1}{s}\left(\exp\left(\frac{R^{min}Q}{N}\right)-1\right) & \text{if } x_{qn}^*=1 \\ 0 & \text{otherwise} \end{cases} \tag{21}$$

where $\left\{x_{qn}^*\right\}_{qn}$ is any spectrum allocation policy such that every band is used and each SU is allocated exactly $\frac{N}{Q}$ bands.

In conclusion, in order to minimize the power consumption of the CR symmetric system, all the spectrum has to be used (see Proposition 1). Since all SUs have the same channel conditions, the spectrum manager does not privilege any particular SU and the spectrum is divided equally among them. There are many ways to allocate the spectrum either by an interleaved or blockwise allocation [35]. These two types of allocations are all equivalent given the symmetry of the system. Then, each SU allocates uniformly its power as in (21) over its allocated bands.

In the following Table 1, we compare the optimal value of the objective function $\sum_q \sum_n p_{qn}^*$ in different particular cases regarding the scheduling policy to illustrate that the fair spectrum allocation among the different users is optimal. We consider the following scenario, which falls into the hypothesis of Proposition 6: $Q = 3$ SUs, $N = 9$ bands, $R^{min} = 3$ bps, T is chosen randomly such that $1 < T < N-2$, we choose $T = 3$ and the SINR $s = 2$. Notice

Table 1 Comparing different spectrum allocations

SU1	SU2	SU3	$\sum_{q,n} p_{qn}^*$ (mW)
$\frac{N}{Q}$	$\frac{N}{Q}$	$\frac{N}{Q}$	2.5774
$\frac{N}{Q}+1$	$\frac{N}{Q}-1$	$\frac{N}{Q}$	3.1585
$\frac{N-1}{2}$	$\frac{N-1}{2}$	1	10.6598
$N-1-T$	T	1	10.8130
$N-2$	1	1	19.3531

that all permutations of $\left\{\frac{N}{Q}, \frac{N}{Q}-1, \frac{N}{Q}+1\right\}$ are equivalent by symmetry and lead to an identical power consumption; therefore, we provide only one of these permutations in the Table. We remind that the sum of the allocated bands must be equal to N i.e. $\sum_q N_q = N$.

We remark that exactly $\frac{N}{Q}$ bands are allocated to each SU. This numerical result validates our Proposition 6.

Next, we provide a more general solution, in the case in which the number of bands is not proportional to the number of SUs, $N/Q \notin \mathbb{N}$. It turns out that the optimal solution to problem (DP_N) is very similar to the fair spectrum allocation discussed above.

Proposition 7. *The optimal solution of the problem (DP_N) when feasible is given as follows:*

$$N_q^* = \begin{cases} \lfloor \frac{N}{Q} \rfloor + 1 & \text{if } q \in \left\{1, \ldots, N - \lfloor \frac{N}{Q} \rfloor Q\right\} \\ \lfloor \frac{N}{Q} \rfloor & \text{if } q \in \{N - \lfloor \frac{N}{Q} \rfloor Q + 1, \ldots, Q\} \end{cases} \quad (22)$$

where $\lfloor y \rfloor$ denotes the integer part of y. The proof of this Proposition is detailed in Appendix D. The optimal power allocation is uniform over all bands that are used by the SUs and is given by (21). The difference with the previous case lies in the number of bands that each SU can use.

In conclusion, when solving the (DP_N) problem in this symmetric system and because of the convex shape of the objective function in a continuous space, two cases arise: (1) either the solution to the relaxed convex continuous problem (i.e., N/Q) is integer and the fair spectrum allocation is optimal $N_q^* = \frac{N}{Q}$, $\forall q$; (2) or the solution $\frac{N}{Q}$ is not an integer and each SU is allocated either $\lfloor \frac{N}{Q} \rfloor$ or $\lfloor \frac{N}{Q} \rfloor + 1$ bands such that $\sum_q N_q^* = N$. This means that the solution is not perfectly fair (as the previous case) and a subset of $\left\{N - \lfloor \frac{N}{Q} \rfloor Q + 1, \ldots, Q\right\}$ SUs will be allocated an additional band.

Aside from this particular case (w.r.t. the system parameters), there are other interesting cases but more complicated to solve: (1) the case in which a *SUq* has the same channel gains and peak interference thresholds over all the frequency bands, but two different SUs do not necessary have the same channel conditions: $s_{qn} = s_q$, $g_{qn} = g_q$, $P_{qn}^{peak} = P_q^{peak}$, $\forall n$; (2) the case in which in a band n all SUs have the same channel gains and peak interference thresholds: $s_{qn} = s_n$, $g_{qn} = g_n$, $P_{qn}^{peak} = P_n^{peak}$, $\forall q$. Although, in the first case (1), we can simplify the problem using the optimal power uniform allocation for each *SUq*, an analytical solution seems difficult or even impossible to find since, as opposed to the perfectly symmetric case, we can no longer conclude from the KKT conditions that the number of allocated bands is the same for all SUs. In the second case, we can no longer use the fact that the optimal power allocation is uniform over the allocated spectrum, so we cannot even simplify the problem as the symmetric case. Since these cases seem difficult and arise with probability zero in practice, we will not detail them here.

6 Conclusions

In this work, we have investigated the usage of full CSI by a cognitive radio manager to jointly schedule spectrum access and power allocation for opportunistic users in a power-efficient CR network. Two main challenges emerge in the underlying optimization problem. A first difficulty lies in the QoS and interference power constraints which may not be met simultaneously. To tackle this issue, we have provided general necessary conditions and sufficient conditions for the existence of an optimal solution.

A second challenge lies in the non-convexity of this problem because of the discrete scheduling policies. This aspect is overcome by exploiting a specific Lagrangian relaxation technique. We have proposed an iterative projected sub-gradient algorithm converging to the optimal joint power and spectrum policy whenever the problem is feasible and the channels are asymmetric. It turns out, that the convergence point is the optimal solution to the initial discrete problem. We have also studied the particular case of a symmetric CR network for which a closed-form solution is found.

Future work may consider further analysis of the cases in which the problem is unfeasible. Instead of not scheduling any SU, the radio-resource manager may decide to remove a subset of SUs chosen (similarly to [8]) to make the optimization problem feasible with a limited outage probability.

Endnotes

[1] We denote

$$[x]_a^b = \begin{cases} a & \text{if } x \leq a \\ x & \text{if } a < x < b \\ b & \text{if } x \geq b \end{cases}$$

[2] We denote by $[x]^+ = \max(x, 0)$.

[3] We denote by $\mathbb{1}_{\{\widetilde{x}_{qn}\}}$ the quantity that equals one if $\widetilde{x}_{qn} = 1$ and zero otherwise.

[4] We denote by $\boldsymbol{s} = [s_{qn}]_{q,n}$ the $Q \times N$ dimensional matrix with entries s_{qn} for all SUs and all bands.

[5] We denote by $\boldsymbol{g}^{(k)} = \left[g_{qn}^{(k)}\right]_{q,n}$ the $Q \times N$ dimensional matrix with entries $g_{qn}^{(k)}$ for all SUs and all bands., $\forall k$, $P_{qn}^{peak(k)} = 20$ mW, $\forall q, \forall n, \forall k$, $R_1^{CS} = 1.4421$ bps, $R_2^{CS} = 4.7488$ bps.

[6] We denote by $Card(\mathcal{X})$ the cardinal value of the set $\mathcal{X}$.

[7] We denote by $||X||_2$ the euclidean norm of X.

Appendix A: Feasibility of the joint scheduling and power allocation problem

In the following, we will provide both necessary conditions and sufficient conditions for the existence of at least one solution to (DP_1) problem.

A1. Proof of Theorem 1, necessary conditions for the existence of a solution

i) We start with a particular case in which $Q = 1$, $N \geq Q, K \geq 1$. There is only one SU occupying all the spectrum *i.e.*, $\forall n, x^*_{1n} = 1$. In this case, we consider the constrained rate maximization problem at the SU level in order to find the maximum achievable rate under interference constraints to PUs. A necessary and sufficient condition (CNS) to have a solution is given by $R_1^{min} \leq R_1^{max}$. Otherwise, if the rate of the SU is too restrictive (i.e., the minimum rate requirement is above R_1^{max}), even if this SU owns all the spectrum, there can be no solution meeting all the constraints and the feasible set is void. We can summarize this necessary and sufficient condition as follows:

$$\begin{cases} Q = 1, N \geq Q, K \geq 1 \\ R_1^{min} \leq R_1^{max} \end{cases} \underset{CNS}{\Leftrightarrow} \mathcal{S}_{\mathcal{F}} \neq \emptyset$$

ii) Now, we assume that $Q \geq 1$, $N \geq Q, K \geq 1$. If we allocate all the spectrum to only one SU, then it must satisfy its rate requirement otherwise the problem cannot have a solution and the feasible set is void. Thus, each SU must have the demand for its minimum rate lower than the rate that it would have if it was the only SU in the spectrum R_ℓ^{max} defined in (5). Assume that it exists an arbitrary SU ℓ, such that it is the only SU in the spectrum i.e., $x_{\ell n} = 1, \forall n$ and if $R_\ell^{min} > R_\ell^{max}$ then there is no solution and $\mathcal{S}_{\mathcal{F}} = \emptyset$. We summarize this necessary condition as follows:

$$\begin{cases} Q \geq 1, N \geq Q, K \geq 1 \\ \exists\, \ell :\ R_\ell^{min} > R_\ell^{max} \end{cases} \underset{CN}{\Rightarrow} \mathcal{S}_{\mathcal{F}} = \emptyset$$

A2. Proof of Theorem 2, sufficient conditions for the existence of a solution

Assume that $Q \geq 1, N \geq Q, \forall K$, we want to find sufficient conditions (CS) on the system parameters that ensure the existence of a pair (x, p) that satisfies all the constraints simultaneously. In order to find these sufficient conditions, we construct $\widetilde{x}$ and $\widetilde{p}$ that satisfy the scheduling and power constraints and are defined by:

$$\widetilde{x}_{qn} = \begin{cases} 1 \ \textit{if} \ \text{q=n} \\ 0 \ \text{otherwise} \end{cases} \tag{23}$$

and

$$\widetilde{p}_{qn} = \min_k \left\{ \frac{\overline{P}_q^{(k)}}{g_{qn}}, \frac{P_{qn}^{peak(k)}}{g_{qn}} \right\} \delta_{q=n}. \tag{24}$$

Such construction of $\widetilde{x}$ is possible because we assumed that $N \geq Q$. Now, we want to find the CS such that the constructed pair $(\widetilde{x}, \widetilde{p})$ (which respects the scheduling and the power constraints) satisfies also the minimum rate constraint. The achievable rate of the SU q under the constructed spectrum and power allocation is

$$\begin{aligned} R_q(\underline{\widetilde{x}}_q, \underline{\widetilde{p}}_q) &= \sum_n \widetilde{x}_{qn} \log_2 \left(1 + \frac{s_{qn}\widetilde{p}_{qn}}{\widetilde{x}_{qn}}\right) \\ &= \widetilde{x}_{qq} \log_2 \left(1 + \frac{s_{qq}\widetilde{p}_{qq}}{\widetilde{x}_{qq}}\right) \\ &= \log_2 \left(1 + s_{qq}\widetilde{p}_{qq}\right). \end{aligned}$$

It turns out that this rate is higher than the rate of the worst channel gain $\min_n\{s_{qn}\}$. Thus, the resulted rate in the previous equality becomes:

$$R_q(\underline{\widetilde{x}}_q, \underline{\widetilde{p}}_q) \geq \log_2 \left(1 + \min_n \left\{ s_{qn} \min_k \left\{ \frac{\overline{P}_q^{(k)}}{g_{qq}}, \frac{P_{qq}^{peak(k)}}{g_{qq}} \right\} \right\}\right). \tag{25}$$

This inequality follows from the definition of the constructed power in (24). So, the rate constraint is always satisfied if the minimum target rate R_q^{min} is lower than this threshold in (25) and then the feasible set is non-void. To summarize, we have the following sufficient conditions:

$$\begin{cases} Q \geq 1, N \geq Q, K \geq 1, \forall\, q \\ R_q^{min} \leq \log_2 \left(1 + \min_n \left\{ s_{qn} \min_k \left\{ \frac{\overline{P}_q^{(k)}}{g_{qq}}, \frac{P_{qq}^{peak(k)}}{g_{qq}} \right\} \right\}\right) \end{cases} \underset{CS}{\Rightarrow} \mathcal{S}_{\mathcal{F}} \neq \emptyset.$$

Remark that tighter sufficient conditions may be obtained directly by using the rates built in (25). They will depend on the system parameters. However, this construction is likely to be unfair from the SUs perspective and rather arbitrary as it is based on the fact that SU q is allocated precisely the band of index q. Any other permutation out of the $Q!$ possibilities may be also considered to obtain better sufficient conditions for some of SUs and conservative for others.

Appendix B: Properties of the relaxed problem

B1. Proof of Proposition 1

We will show that rate constraint is satisfied with equality at the optimum:

$$\sum_n x^*_{qn} \log_2 \left(1 + \frac{p^*_{qn} s_{qn}}{x^*_{qn}}\right) = R_q^{min}, \forall q.$$

Assume on the contrary (by *Reductio ad absurdum*) that there exists a SU q such that at the optimum $(\boldsymbol{x}^*, \boldsymbol{p}^*)$: $\sum_m x^*_{qm} \log_2 \left(1 + \frac{p^*_{qm} s_{qm}}{x^*_{qm}}\right) > R_q^{min}$. We choose an arbitrary band n and we prove that, in this case, SU q can reduce its overall power consumption while still meeting the constraints which leads to a contradiction. We have

$$\sum_{m\neq n} x^*_{qm}\log_2\left(1+\frac{s_{qm}p^*_{qm}}{x^*_{qm}}\right)+x^*_{qn}\log_2\left(1+\frac{s_{qn}p^*_{qn}}{x^*_{qn}}\right)> R^{min}_q$$

equivalent to

$$\log_2\left(1+s_{qn}p^*_{qn}\right) > R^{min}_q-\sum_{m\neq n} r^*_{qm},$$

where $r^*_{qm}=x^*_{qm}\log_2\left(1+\frac{p^*_{qm}s_{qm}}{x^*_{qm}}\right)$.

We build a new allocation $(\boldsymbol{x}^{new},\boldsymbol{p}^{new})$ as follows: the scheduling allocation policy is the same as the optimum one $\boldsymbol{x}^{new}=\boldsymbol{x}^*$ and the power allocation policy is the same as the optimum one $p^{new}_{qm}=p^*_{qm}$ except in the n band in which we decrease power in order to achieve the rate R^{min}_q with equality such that:

$$p^{new}_{qm}=\begin{cases} p^*_{qm},\ \ if\ \ m\neq n \\ p^{new}_{qn},\ if\ \ m=n \end{cases},$$

where $p^{new}_{qn}=\frac{1}{s_{qn}}\left(\exp\left(R^{min}_q-\sum_{m\neq n} r^*_{qm}\right)-1\right)$ and we have: $p^{new}_{qn}<p^*_{qn}$. Thus, the objective function of power consumption: $\sum_n p^{new}_{qn}<\sum_n p^*_{qn}$. With this new power allocation, the user q satisfies all the constraints and has a strictly lower power consumption than the optimal power p^* which is a contradiction. So, we conclude that $\forall\ q$, $\sum_n x^*_{qn}\log_2\left(1+\frac{p^*_{qn}s_{qn}}{x^*_{qn}}\right)=R^{min}_q$.

B2. Proof of Proposition 2

Now, we will show that scheduling constraints are satisfied also with equality at the optimal power and spectrum allocation: $\forall\ n$, $\sum_q x^*_{qn}=1$.

Assume on the contrary (by *Reductio ad absurdum*) that there exists a band n such that $\sum_q x^*_{qn}<1$. We build another allocation $(\boldsymbol{x}^{new},\boldsymbol{p}^{new})$ as follows: the allocation of the spectrum is the same as the optimum one $\boldsymbol{x}^{new}=\boldsymbol{x}^*$ except for the band n which we decide to allocate to an arbitrary user k as follows:

$$x^{new}_{km}=\begin{cases} x^*_{km} & if\ \ m\neq n \\ 1-\sum_q x^*_{qn} & if\ \ m=n. \end{cases} \tag{26}$$

Therefore, we have that $x^{new}_{kn}>x^*_{kn}$. Since we know, from Proposition 1, that the rate constraint is met with equality at the optimal solution, we build the new power allocation vector $\boldsymbol{p}^{new}$ such that the allocation of the power is the same as the optimum one $\boldsymbol{p}^{new}=\boldsymbol{p}^*$ except for the user k for which: $\sum_m x^{new}_{km}\log_2\left(1+\frac{p^{new}_{km}s_{km}}{x^{new}_{km}}\right)=R^{min}_k$ which is equivalent to: $\sum_{m\neq n} x^*_{km}\log_2\left(1+\frac{p^*_{km}s_{km}}{x^*_{km}}\right)+\left(1-\sum_q x^*_{qn}\right)\log_2\left(1+\frac{p^{new}_{kn}s_{kn}}{1-\sum_q x^*_{qn}}\right)=R^{min}_k$.

Then, we obtain:

$$p^{new}_{kn}=\frac{1-\sum_q x^*_{qn}}{s_{kn}}\left(\exp\left(\frac{R^{min}_k-\sum_{m\neq n} r^*_{km}}{1-\sum_q x^*_{qn}}\right)-1\right).$$

However, the optimal power allocation is given by

$$p^*_{kn}=\frac{x^*_{kn}}{s_{kn}}\left(\exp\left(\frac{R^{min}_k-\sum_{m\neq n} r^*_{km}}{x^*_{kn}}\right)-1\right).$$

Knowing that $x^*_{kn}<1-\sum_q x^*_{qn}$ and that the function $f(X)=\frac{X}{S}\left(\exp\left(\frac{R}{X}\right)-1\right)$ is decreasing, we obtain that $p^*_{kn}>p^{new}_{kn}$. Therefore, user k achieves the rate R^{min}_k with lower consumption than the optimal point $(\boldsymbol{x}^*,\boldsymbol{p}^*)$

$$\sum_m p^{new}_{km}<\sum_m p^*_{km}.$$

Thus, we have a pair $(\boldsymbol{x}^{new},\boldsymbol{p}^{new})$ that satisfies all the constraints and gives us a strictly lower power consumption than the optimum which is a contradiction. Then, all average scheduling constraints are satisfied with equality.

In conclusion, if a SU uses an extra bandwidth, it achieves the same rate R^{min}_k with less power consumption. Therefore, at the optimum all bands are entirely used.

Appendix C: Convergence proof

In the following, we will prove that if the feasible set of problem (CP_2) is non-void, then our projected sub-gradient Algorithm 1 converges always to an optimal solution.

In order to prove the convergence of our algorithm, we start by proving that our dual function $\mathrm{G}(\boldsymbol{\lambda},\boldsymbol{\beta})$ in **Step 4** is concave with respect to $(\boldsymbol{\lambda},\boldsymbol{\beta})$.

Given that the Lagrangian $\mathrm{L}(\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\mu},\mathbf{x},\mathbf{p})$ in **Step 1** is affine in $\boldsymbol{\lambda},\boldsymbol{\beta}$ and $\boldsymbol{\mu}$ for any feasible $(\mathbf{x},\ \mathbf{p})$ and according to [14], the point-wise infimum $\mathrm{G}(\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\mu})=\min_{(\mathbf{x},\mathbf{p})}\mathrm{L}(\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\mu},\mathbf{x},\mathbf{p})$ is then jointly concave w.r.t $\boldsymbol{\lambda},\boldsymbol{\beta}$ and $\boldsymbol{\mu}$. Since the resulting dual function $\mathrm{G}(\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\mu})$ is jointly concave w.r.t $(\boldsymbol{\lambda},\boldsymbol{\beta})$ and $\boldsymbol{\mu}$ and according to the theorem detailed in [36] (Appendix B.15), we obtain the concavity of the function $\mathrm{G}(\boldsymbol{\lambda},\boldsymbol{\beta})=\max_{\boldsymbol{\mu}}\mathrm{G}(\boldsymbol{\lambda},\boldsymbol{\beta},\boldsymbol{\mu})$.

Inspired by [14] and since $\mathrm{G}(\boldsymbol{\lambda},\boldsymbol{\beta})=-\sum_n\max_q\varphi_{qn}(\lambda^{(k)}_q,\beta_q)+\sum_q\beta_q R^{min}_q-\sum_k\sum_q\lambda^{(k)}_q\overline{P}^{(k)}_q$ is jointly concave w.r.t $\boldsymbol{\lambda}$ and $\boldsymbol{\beta}$ but not differentiable because of the piecewise function φ_{qn}, in order to maximize $\mathrm{G}(\boldsymbol{\lambda},\boldsymbol{\beta})$, we propose to use a projected sub-gradient method such that

$$X^{(t+1)}=\left[X^{(t)}-\tau D^{(t)}\right]^+, \tag{27}$$

where $X^{(t)}$ is the update of the problem variable $X=\begin{pmatrix}\boldsymbol{\lambda}\\ \boldsymbol{\beta}\end{pmatrix}$ at time t. We project this variable because of the positivity

of $\boldsymbol{\lambda}$ and $\boldsymbol{\beta}$ and $D^{(t)}$ is any sub-gradient of G at $X^{(t)}$ and $\tau > 0$ is a constant step size.

We choose the sub-gradient of the dual function $\mathrm{G}(\boldsymbol{\lambda}, \boldsymbol{\beta})$ at $\boldsymbol{\lambda}, \boldsymbol{\beta}$ given by the constraints

$$D(\boldsymbol{\lambda}, \boldsymbol{\beta}) = \begin{pmatrix} -\sum\limits_{n} g_{qn}^{(k)} p_{qn} + \overline{P}_q^{(k)} \\ R_q(\underline{x}_q, \underline{p}_q) - R_q^{min} \end{pmatrix}$$

The proof of this proposition is based on the fact that a sub-gradient of the function G at $(\boldsymbol{\lambda}, \boldsymbol{\beta})$ is any vector D that satisfies the inequality, $\forall(\boldsymbol{\lambda}_1, \boldsymbol{\beta}_1)$ and $(\boldsymbol{\lambda}_2, \boldsymbol{\beta}_2)$:

$$\mathrm{G}(\boldsymbol{\lambda}_1, \boldsymbol{\beta}_1) - \mathrm{G}(\boldsymbol{\lambda}_2, \boldsymbol{\beta}_2) \leq D(\boldsymbol{\lambda}_2, \boldsymbol{\beta}_2)^T \begin{pmatrix} \boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2 \\ \boldsymbol{\beta}_1 - \boldsymbol{\beta}_2 \end{pmatrix}. \quad (28)$$

Since $\boldsymbol{x}$ and $\boldsymbol{p}$ are bounded such that $0 \leq g_{qn}^{(k)} p_{qn} \leq P_{qn}^{peak(k)}$ and $0 \leq x_{qn} \leq 1$, thus, the norm[7] of the sub-gradient is also bounded $||D(\boldsymbol{\lambda}, \boldsymbol{\beta})^{(t)}||_2 \leq D_{th}$ such that

$$D_{th} = \sqrt{D_{th1}^2 + D_{th2}^2} \text{ with}$$

$$D_{th1} = \max\left\{ \min_k \left\{ \overline{P}_q^{(k)} \right\}, \min_k \left\{ \overline{P}_q^{(k)} \right\} - N \min_k \left\{ P_{qn}^{peak} \right\} \right\},$$

$$D_{th2} = \sum_n \log_2 \left(1 + s_{qn} \min_k \left\{ \frac{P_{qn}^{peak}}{g_{qn}} \right\} \right) - R_q^{min}.$$

From [14], we have that our sub-gradient algorithm converges, when $t \to +\infty$, to a neighborhood of the optimal solution depending on the step-size:

$$\max_{i=1,\ldots,t} ||G\left(\boldsymbol{\lambda}^{(i)}, \boldsymbol{\beta}^{(i)}\right) - G(\boldsymbol{\lambda}^*, \boldsymbol{\beta}^*)|| \leq \frac{D_{th}^2 \tau}{2}.$$

Regarding such sub-gradient methods, there are a lot of convergence results (such as point-wise convergence) available [37, 38] for different choices of the step-size other than constant: constant step length; square summable but not summable (e.g., $\tau^{(t)} = a/(b+t)$, $a > 0$ and $b \geq 0$); and non-summable diminishing (e.g., $\tau^{(t)} = a/\sqrt{t}$, $a > 0$). The constant step-size and constant step-length choices guarantee that the sub-gradient method converges to a certain neighborhood of the optimal solution set, where the size of that neighborhood depends on the value of the step-size, in our case τ. But, if we want to obtain stronger convergence results, we will have to use variable step-sizes [14, 38].

Appendix D: Extreme case where channel gains are uniform for each SU and over each band

We are interested in a particular symmetric case which cannot be solved by our Algorithm 1. In such a case, there are ties regarding the spectrum allocation as we can schedule one band to several SUs with no optimality loss. In this case, we assume that $s_{qn} = s$, $g_{qn} = g$, $P_{qn}^{peak} = P^{peak}$, $\overline{P}_q = \overline{P}$, $R_q^{min} = R^{min}$, $\forall n, \forall q$. In the following, we will prove Propositions 5, 6 and 7.

Proof of Proposition 6 and Proposition 7

First, we start by solving a simpler problem than (DP_N) by ignoring the interference constraints $(C1)$ and $(C2)$ and taking into account only the scheduling constraint $(C3)$. We begin by studying a relaxed version of this problem in which N_q is a positive real:

$$\begin{aligned} &\text{minimize} \sum_{q \in \mathcal{Q}} \frac{N_q}{s} \left(\exp\left(\frac{R^{min}}{N_q} \right) - 1 \right) \\ &\text{subject to} \sum_{q=1}^{Q} N_q = N \\ &\qquad N_q > 0, \ \forall q. \end{aligned}$$

This problem is convex because its the objective function is jointly convex in $(N_1, N_2, \ldots, N_Q)$ and the constraints are affine in N_q. The Lagrangian of this problem is given by:

$$\begin{aligned} \mathrm{L}(\boldsymbol{\alpha}, \boldsymbol{\delta}, \mathbf{N}_q) = &\sum_{q \in \mathcal{Q}} \frac{N_q}{s} \left(\exp\left(\frac{R^{min}}{N_q} \right) - 1 \right) \\ &+ \alpha \left(\sum_{q=1}^{Q} N_q - N \right) - \sum_q \delta_q N_q \end{aligned}$$

The KKT optimality conditions for this continuous problem are:

1) $\dfrac{\partial \mathrm{L}(\boldsymbol{\alpha}, \boldsymbol{\delta}, \mathbf{N}_q)}{\partial N_q^*} = 0, \ \forall q$

$\Rightarrow \exp\left(\dfrac{R^{min}}{N_q^*} \right) \left(\dfrac{1}{s} - \dfrac{R^{min}}{s N_q^*} \right) - \dfrac{1}{s} + \alpha^* - \delta_q^* = 0$

2) $\alpha^* \in \mathbb{R}$ and $\sum\limits_q N_q^* = N$

3) $\forall q, \delta_q^* = 0$ and $N_q^* > 0$

From all these KKT conditions, we obtain the system of equations:

$$\begin{cases} \exp\left(\dfrac{R^{min}}{N_q^*} \right) \left(\dfrac{1}{s} - \dfrac{R^{min}}{s N_q^*} \right) - \dfrac{1}{s} + \alpha^* = 0, \ \forall q \\ \alpha^* \in \mathbb{R} \ \text{ and } \ \sum\limits_q N_q^* = N \\ N_q^* > 0, \ \forall q. \end{cases}$$

From the first equation, we can write:

$$\alpha^* = -\exp\left(\frac{R^{min}}{N_q^*} \right) \left(\frac{1}{s} - \frac{R^{min}}{s N_q^*} \right) + \frac{1}{s}, \ \forall q \quad (29)$$

From Eq. (29) and the fact that the function $U(M) = \frac{1}{s} - \frac{1}{s}\exp\left(\frac{R^{min}}{M}\right)\left(1 - \frac{R^{min}}{M}\right)$ is strictly increasing and bijective, we conclude that at the optimum, we have the same number of allocated bands $N_q^* = N_r^*,\ \forall q \neq r$. Considering the constraint $\sum_q N_q^* = N$ and using the fact that the objective function $\sum_q f(N_q)$ is convex, we have two cases for the solution of the discrete problem (DP_N) without the constraints $(C1)$ and $(C2)$:

a) If the minimum of the objective function of the convex continuous problem $\frac{N}{Q}$ is an integer then N_q^* is uniform for all the SUs $N_q^* = \frac{N}{Q}$, for all SUs.

b) If the minimum of the objective function $\frac{N}{Q}$ is not an integer $\frac{N}{Q} \notin \mathbb{N}$ then a SUq is allocated either $\lfloor \frac{N}{Q} \rfloor$ or $\lfloor \frac{N}{Q} \rfloor + 1$ bands such that $\sum_q N_q^* = N$. This means that the solution is not perfectly fair as the previous case and a subset of $\left\{N - \lfloor \frac{N}{Q} \rfloor Q + 1, \ldots, Q\right\}$ SUs will be allocated one band more:

$$N_q^* = \begin{cases} \lfloor \frac{N}{Q} \rfloor + 1 & \text{if } q \in \{1, \ldots N - \lfloor \frac{N}{Q} \rfloor Q\} \\ \lfloor \frac{N}{Q} \rfloor & \text{if } q \in \{N - \lfloor \frac{N}{Q} \rfloor Q + 1, \ldots, Q\} \end{cases} \quad (30)$$

Now, the objective is to prove that if (DP_N) is feasible, then the optimal solution is either $N_q^* = \frac{N}{Q}$, for all SUs if $\frac{N}{Q}$ is an integer or N_q^* in (30) if $\frac{N}{Q}$ is not an integer. From the feasibility assumption, there exists at least one feasible $(\widetilde{N}_1, \ldots, \widetilde{N}_Q)$ that satisfies $(C1)$, $(C2)$, $(C3)$, and $(C4)$. From condition $(C3)$, we have that there exists at least one index q such that $\widetilde{N}_q \leq \lfloor N/Q \rfloor$ (otherwise $(C3)$ is not met). So, we have $N/Q \geq \lfloor N/Q \rfloor \geq \widetilde{N}_q$.

First, from condition $(C1)$ we have $f(\widetilde{N}_q) \leq \frac{\overline{P}}{g}$. The fact that $f(M) = M/s*(exp(R^{min}/M)-1)$ is strictly decreasing in M implies

$$f(N/Q) \leq f(\lfloor N/Q \rfloor) \leq f(\widetilde{N}_q) \leq \frac{\overline{P}}{g}.$$

Also, we have the trivial inequality $\lfloor N/Q \rfloor + 1 \geq \lfloor N/Q \rfloor$ which leads to

$$f(\lfloor N/Q \rfloor + 1) \leq f(\lfloor N/Q \rfloor) \leq \frac{\overline{P}}{g}.$$

Second, from condition $(C2)$ we have: $T(\widetilde{N}_q) \leq \frac{P^{peak}}{g},\ \forall q$ where $T(M) = 1/s * (exp(R^{min}/M) - 1)$ is the peak power required to reach R^{min} with uniform allocation on M bands. We know that $T(M)$ is strictly decreasing in M, which implies

$$T(N/Q) \leq T(\lfloor N/Q \rfloor) \leq T(\widetilde{N}_q) \leq \frac{P^{peak}}{g},$$

and also $T(\lfloor N/Q \rfloor + 1) \leq T(\lfloor N/Q \rfloor) \leq \frac{P^{peak}}{g}$.

In conclusion, if (DP_N) is feasible, then the optimal solution is either $N_q^* = \frac{N}{Q}$ (for all SUs) provided $\frac{N}{Q}$ is an integer, or N_q^* in (30) otherwise.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]ETIS/ ENSEA, University Cergy Pontoise, CNRS, F-95000 Cergy Pontoise, France. [2]LETI, ENIS and University of Sfax, Sfax, Tunisia. [3]INRIA Grenoble Rhône-Alpes, Grenoble, France.

Received: 21 October 2015 Accepted: 28 March 2016
Published online: 11 April 2016

References

1. Spectrum Policy Task Force, Report of the spectrum rights and responsibilities working group, tech. rep., Federal Communications Commission (2002)
2. JM Peha, Sharing spectrum through spectrum policy reform and cognitive radio. Proc. IEEE. **97**(4), 708–719 (2009)
3. G Scutari, DP Palomar, MIMO cognitive radio: a game theoretical approach. IEEE Trans. Signal Process. **58**(2), 761–780 (2010)
4. D Lee, *et al*, Coordinated multipoint transmission and reception in LTE-advanced: deployment scenarios and opertaional challenges. IEEE Commun. Mag. **50**(2), 148–155 (2012)
5. KI Pedersen, F Frederiksen, C Rosa, H Nguyen, LGU Garcia, Y Wang, Carrier aggregation for LTE-advanced: functionality and performance aspects. IEEE Commun. Mag. **49**(6), 89–95 (2011)
6. G J-Sh Pang, F Scutari, Ch Facchinei, Wang, Distributed power allocation with rate constraints in gaussian parallel interference channels. IEEE Trans. Inf. Theory. **54**(8), 3471–3489 (2008)
7. R Cavalcante, *et al*, Toward energy-efficient 5G wireless communications technologies: tools for decoupling the scaling of networks from the growth of operating power. IEEE Signal Proc. Mag. **31**(6), 24–34 (2014)
8. A Alsawah, I Fijalkow, Practical radio link resource allocation for fair QoS-provision on OFDMA downlink with partial channel-state information. EURASIP J. Adv. Signal Process. **2009**(1), 1–16 (2008)
9. D Gesbert, MS Alouini. Selective multi-user diversity, Internaltional Symposium on Signal Processing and Information Technology, (Germany, 2003), pp. 162–165
10. DT Ngo, T Le-Ngoc, Distributed resource allocation for cognitive radio networks with spectrum-sharing constraints. IEEE Trans. Veh. Technol. **60**(7), 3436–3449 (2011)
11. G Scutari, DP Palomar, S Barbarossa, Asynchronous iterative water- filling for Gaussian frequency-selective interference channels. IEEE Trans. Inf. Theory. **54**(7), 2868–2878 (2008)
12. D Palomar, J Fonollosa, Practical algorithms for a family of waterfilling solutions. IEEE Trans. Signal Process. **53**(2), 686–695 (2005)
13. J Huang, VG Subramanian, R Agrawal, RA Berry, Joint scheduling and resource allocation in uplink OFDM systems for broadband wireless access networks. IEEE JSAC. **27**(2), 226–234 (2009)
14. S Boyd, L Xiao, A Mutapcic, *Subgradient Methods*, vol. 2004. (Notes for EE392o, Stanford University, Autumn Quarter, 2003), pp. 2004–2005
15. DD Yu, Cioffi JM, Spc10-2: iterative water-filling for optimal resource allocation in ofdm multiple-access and broadcast channels. IEEE Glob. Telecommun. Conf (2006)
16. G Scutari, DP Palomar, S Barbarossa, Optimal linear precoding strategies for wideband noncooperative systems based on game theory–part I: Nash equilibria. IEEE Trans. Signal Process. **56**(3), 1230–1249 (2008)
17. G Scutari, DP Palomar, S Barbarossa, Optimal linear precoding strategies for wideband non-cooperative systems based on game theory–part II: algorithms. IEEE Trans. Signal Process. **56**(3), 1250–1267 (2008)
18. EV Belmega, S Lasaulce, M Debbah, Power allocation games for MIMO multiple access channels with coordination. IEEE Trans. Wirel. Commun. **8**(6), 3182–3192 (2009)
19. G Miao, N Himayat, YG Li, Crosslayer optimization for energy efficient wireless communications: a survey. Wirel. Commun. Mob. Comput. **9**(4), 529–542 (2009)

20. Ch Isheden, Z Chong, E Jorswieck, Framework for link-level energy efficiency optimization with informed transmitter. IEEE Trans. Wirel. Commun. **11**(8), 2946–2957 (2012)
21. G Miao, N Himayat, YG Li, Distributed interference-aware energy-efficient power optimization. IEEE IEEE Trans. Wirel. Commun. **10**(4), 1323–1333 (2011)
22. G Bacci, EV Belmega, L Sanguinetti, Distributed energy-efficient power optimization in cellular relay networks with minimum rate constraints. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP), 7014–7018 (2014)
23. VG Subramanian, RA Berry, R Agrawal, Joint scheduling and resource allocation in CDMA systems. IEEE Trans. Inf. Theory. **56**(5), 2416–2432 (2010)
24. J Huang, VG Subramanian, R Agrawal, RA Berry, Downlink scheduling and resource allocation for OFDM systems. IEEE Trans. Wirel. Commun. **8**(1), 288–296 (2009)
25. A Bagayoko, I Fijalkow, P Tortelier, Power control of spectrum-sharing in fading environment with partial channel state information. IEEE Trans. Signal Process. **59**(5), 2244–2256 (2011)
26. P Mertikopoulos, EV Belmega, Transmit without regrets: online optimization in MIMO-OFDM cognitive radio systems. IEEE J. Sel. Areas Commun. **32**(11), 1987–1999 (2014)
27. R Masmoudi, EV Belmega, I Fijalkow, N Sellami, in *International Symposium DySPAN*. A closed-form solution to the power minimization problem over two orthogonal frequency bands under QoS and cognitive radio interference constraints, pp. 212–222
28. R Masmoudi, EV Belmega, I Fijalkow, N Sellami, A unifying view on energy-efficiency metrics in cognitive radio channels. Eur. Signal Process. Conf, 171–175 (2014)
29. SN Diggavi, TM Cover, The worst additive noise under a covariance constraint. IEEE Trans. Inf. Theory. **47**(7), 3072–3081 (2001)
30. E Telatar, Capacity of multi-antenna Gaussian channels.European Trans. Telecommun. Wiley Online Libr. **10**(6), 585–595 (1999)
31. CE Shannon, Communication in the presence of noise. Proc IRE. **37**(1), 10–21 (1949)
32. E Masera, T Lestable, *Econhome project presentation*. (Mobile World Congress, Barcelona, 2011)
33. H Congzheng, T Harrold, *et al*, Green radio: radio techniques to enable energy-efficient wireless networks. IEEE Commun. Mag. **49**(6), 46–54 (2011)
34. S Boyd, L Vandenberghe, *Convex Optimization*. (Cambridge Univ. Press, Cambridge, 2004)
35. AM Masucci, EV Belmega, I Fijalkow, *Optimal blockwise subcarrier allocation policies in single carrier FDMA uplink systems*, (2014)
36. S Foucart, H Rauhut, *A mathematical introduction to compressive sensing*. (Birkhäuser, Boston, 2013)
37. Nesterov Y, Introductory Lectures on Convex Programming Volume I: Basic course, Lecture notes (1998)
38. NZ Shor, *Minimization methods for non-differentiable functions*, vol. 3. (Springer-Science & Business Media, 2012)

Energy-Aware Competitive Power Allocation for Heterogeneous Networks Under QoS Constraints

Giacomo Bacci, *Member, IEEE*, E. Veronica Belmega, *Member, IEEE*, Panayotis Mertikopoulos, *Member, IEEE*, and Luca Sanguinetti, *Member, IEEE*

***Abstract*—This work proposes a distributed power allocation scheme for maximizing energy efficiency in the uplink of OFDMA-based HetNets where a macro-tier is augmented with small cell access points. Each user equipment (UE) in the network is modeled as a rational agent that engages in a non-cooperative game and allocates its available transmit power over the set of assigned subcarriers to maximize its individual utility (defined as the user's throughput per Watt of transmit power) subject to a target rate requirement. In this framework, the relevant solution concept is that of Debreu equilibrium, a generalization of the concept of Nash equilibrium. Using techniques from fractional programming, we provide a characterization of equilibrial power allocation profiles. In particular, Debreu equilibria are found to be the fixed points of a water-filling best response operator whose water level is a function of rate constraints and circuit power. Moreover, we also describe a set of sufficient conditions for the existence and uniqueness of Debreu equilibria exploiting the contraction properties of the best response operator. This analysis provides the necessary tools to derive a power allocation scheme that steers the network to equilibrium in an iterative and distributed manner without the need for any centralized processing. Numerical simulations are used to validate the analysis and assess the performance of the proposed algorithm as a function of the system parameters.**

***Index Terms*—Heterogeneous networks, 5G, energy efficiency, area spectral efficiency, power allocation, distributed algorithms, game theory, Debreu equilibrium, rate constraints.**

I. Introduction

OWING to the prolific spread of Internet-enabled mobile devices and the ever-growing volume of mobile communication calls, the biggest challenge in the wireless industry today is to meet the soaring demand for wireless broadband required to ensure consistent quality of service (QoS) in a network. Rising to this challenge means increasing the network capacity by a thousandfold over the next few years [1], but the resulting power consumption and energy-related pollution are expected to give rise to major societal, economic and environmental issues that would render this growth unsustainable [2]. Therefore, the information and communications technology (ICT) industry is faced with a formidable mission: cellular network capacity must be increased significantly in order to accommodate higher data rates, but this task must be accomplished under an extremely tight energy budget.

A promising way out of this gridlock is the small-cell (SC) network paradigm which builds on the premise of shrinking wireless cell sizes in order to bring user equipment (UE) and their serving stations closer to one another. From an operational standpoint, SC networks can be integrated into existing macrocellular networks: the latter ensure wide-area coverage and mobility support, while the former carry most of the generated data traffic [3].

Albeit promising, the deployment of this kind of networks, commonly referred to as heterogeneous networks (HetNets), poses several technical challenges mainly because different SCs are likely to be connected over unreliable infrastructures with widely varying features, such as error rate, outage, delay, and/or capacity specifications. Accordingly, the inherently heterogeneous nature of these networks calls for flexible and decentralized resource allocation strategies that rely only on local channel state information (CSI) and require minimal information exchange between network users and/or access points/base stations. This framework is commonly referred to as *distributed optimization*, and it represents a crucial aspect of scalable and efficient network operation.

An established theoretical tool for problems of this kind is provided by the theory of *non-cooperative games* [4]. Among the early contributions in this area, [5], [6] investigated the rate maximization problem for autonomous digital subscriber lines based on competitive optimality criteria. In the spirit of these works, a vast corpus of literature has since focused on developing power control techniques for unilateral spectral efficiency maximization subject to individual power constraints. For instance, [7], [8] proposed a game-theoretic approach to energy-efficient power control in multi-carrier code division multiple access (CDMA) systems, [9]–[12] investigated the problem of distributed power control in multi-user multiple-input and multiple-output (MIMO) systems, [13], [14] studied the interference relay channel, while two-tier CDMA networks

Manuscript received November 8, 2014; revised February 6, 2015; accepted April 8, 2015. Date of publication April 22, 2015; date of current version September 7, 2015. The research of G. Bacci and L. Sanguinetti has been funded by the People Programme FP7 PIOF-GA-2011-302520 GRAND-CRU and PIEF-GA-2012-330731 Dense4Green. P. Mertikopoulos was also supported by the CNRS under grant no. PEPS-GATHERING-2014. This research is also supported by NEWCOM# (Grant agreement no. 318306), and by French National Research Agency (ANR) research grant NETLEARN (contract no. ANR-13-INFR-004). Part of this work has been presented at the IEEE International Conference on Communications, Sydney, Australia, June 2014. The associate editor coordinating the review of this paper and approving it for publication was M. Bennis.

G. Bacci is with Mediterranean Broadband Infrastructure (MBI) srl, Pisa 56121, Italy (e-mail: gbacci@mbigroup.it).

E. V. Belmega is with the ETIS/ENSEA, Université de Cergy-Pontoise, Cergy-Pontoise 95014, France (e-mail: elena-veronica.belmega@ensea.fr).

P. Mertikopoulos is with the French National Center for Scientific Research (CNRS), and with the Laboratoire d'Informatique de Grenoble, Grenoble 38330, France (e-mail: panayotis.mertikopoulos@imag.fr).

L. Sanguinetti is with the Dipartimento di Ingegneria dell'Informazione, University of Pisa, Pisa 56122, Italy, and also with Large Networks and Systems Group, CentraleSupélec, Gif-sur-Yvette 91192, France (e-mail: luca.sanguinetti@unipi.it).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2015.2425397

were examined in [15]. More recently, the authors of [16] used a variational inequality (VI) framework to model and analyze the competitive spectral efficiency maximization problem. The analogy between Nash equilibria and VIs was subsequently exploited in [17] to design distributed power control algorithms for spectral efficiency maximization under interference temperature constraints in a cognitive radio context. A possible application of the VI theory for power control in networks with heterogenous UEs can be also found in [18].

Distributed power allocation policies as above have the important advantage of avoiding the waste of energy associated with centralized algorithms requiring considerable information exchange (and, hence, transmissions) between the users and/or the network administrator [16]. On the other hand, the users' aggressive attitude towards interference from other users can lead to a cascade of power increases at the UE level, thereby leading to battery depletion and inefficient energy use. Consequently, solutions that focus exclusively on spectral efficiency maximization are not aligned with energy-efficiency requirements [19], [20], which, as we mentioned above, are crucial for the deployment and operation of HetNets.

A. Summary of Contributions

Our main goal in this paper is the analysis and design of energy-efficient power allocation policies in a HetNet setting where SC networks coexist with macro-tier cellular systems based on orthogonal frequency-division multiple access (OFDMA) technology. In particular, focusing on the uplink case, we propose a game-theoretic framework where each UE adjusts the allocation of its transmit power (over the available subcarriers) so as to unilaterally maximize its individual link utility subject to a minimum rate requirement. Specifically, each user's energy-aware utility function is defined as the achieved throughput per unit power, accounting for both the power required for data transmission and that required by the circuit components of each UE (such as amplifiers, mixeroscillator, and filters) [21]–[23].

Due to each user's rate constraint, the resulting game departs from the classical framework put forth by Nash [24] and gives rise to a Debreu-type game [25] where the actions available to each UE depend on the transmit power profile of all other users in the network. In this setting, the relevant solution concept is that of the *Debreu equilibrium* (DE) [25]—also known as a generalized Nash equilibrium (GNE) [26]. Drawing on fractional programming techniques [27], we characterize the system's Debreu equilibria as fixed points of a water-filling operator whose water level is a function of the users' minimum rate constraints and circuit power [23]. This characterization is then used to provide sufficient conditions for DE uniqueness and to derive a distributed power allocation algorithm that allows the network to converge to equilibrium under minimal information assumptions. The performance of the proposed solution is then validated by means of extensive numerical simulations modeling a HetNets where a macro-tier is augmented with a certain number of low range small-cell access points (SCAs). As turns out, the proposed solution represents a scalable and flexible technique to meet the ambitious goals of 5G communications [28], such as extremely high area spectral efficiency (ASE) (more than 500 b/s/Hz/km^2) with a reasonable amount of physical resources (bandwidth and power) and complexity at the network level (number of SCs, signal processing burden, and number of transmit and receive antennas).

Our work builds on the game-theoretic analysis proposed in [29] where a group of players aims at maximizing their individual energy efficiency (EE) (measured in bits per Watt of transmit power) subject to each user's power constraint. Despite this similarity, the analysis of [29] does not account for minimum rate requirements, thus the resulting game-theoretic model is a standard Nash game with no QoS guarantees–in particular, the users' rates at equilibrium could be fairly low. Incorporating QoS requirements changes the setting drastically and takes us beyond the standard Nash framework because a user's admissible power allocation policy depends crucially on the transmit powers of all other users. The energy-efficient framework proposed in this paper represents a generalization of the power minimization under minimum-rate constraints investigated in [30], which is a special case that occurs when the minimum rates are achieved with equality. Preliminary versions of our results appeared in the conference paper [31]: in contrast to this earlier paper, we provide here a complete equilibrium analysis and characterization along with sufficient conditions that guarantee the convergence of the system to a stable equilibrium state.

B. Paper Outline and Notation

The remainder of this paper is organized as follows. In Section II, we introduce the system model and the EE maximization problem with minimum rate constraints. In Section III, we first formulate the non-cooperative game and then study the existence and uniqueness of Debreu equilibria. Section IV presents an iterative and distributed algorithm to reach the equilibrium point, whereas Section V reports numerical results that are used to assess the performance of the proposed solution and to make comparisons with alternatives. Conclusions and perspectives are presented in Section VI.

Matrices and vectors are denoted by bold letters, $\mathbf{I}_L$, $\mathbf{0}_L$, and $\mathbf{1}_L$ are the $L\times L$ identity matrix, the $L\times 1$ all-zero column vector, and the $L\times 1$ all-one column vector, respectively, and $\|\cdot\|$, $(\cdot)^T$ and $(\cdot)^H$ denote Euclidean norm of the enclosed vector, transposition and Hermitian conjugation respectively. The notation $(x)^+$ stands for $\max\{0,x\}$ whereas $W(\cdot)$ denotes the Lambert W function [32], defined as the multiple-branch solution of the equation $z=W(z)e^{W(z)}$, $z\in\mathbb{C}$. $\mathbb{1}_X$ denotes the indicator function such that $\mathbb{1}_X=1$ if X is true, and 0 elsewhere. Finally, if $\mathcal{A}_k$, $k=1,\ldots,K$, is a finite family of sets, and $a_k\in\mathcal{A}_k$, we will use the notation $(a_k;\mathbf{a}_{-k})\in\prod_k\mathcal{A}_k$ as shorthand for the profile $(a_1,\ldots,a_k,\ldots,a_K)$, and $|\mathcal{A}_k|$ to denote its cardinality.

II. System Model and Problem Formulation

A. System Model

We consider the uplink of a slowly-varying HetNet where S low-range SCA are adjoined to a macro-tier cell operating in an

OFDMA-based open-access licensed spectrum. For notational compactness, we will reserve the index $s = 0$ for the macrocell base station (MBS), so that $\mathcal{S} = \{0, 1, \ldots, S\}$ represents the set of HetNet receiving stations. The s-th cell uses a set of orthogonal subcarriers to serve the K_s user equipment (UE) falling within its coverage radius ρ_s. For simplicity, we assume that the same set of subcarriers $\mathcal{N} = \{1, \ldots, N\}$ is used by both tiers. We also assume that $\mathcal{N}$ is assigned by the network and cannot be controlled by the cell operators. Each cell access point (AP) is further equipped with M_s receiving antennas, whereas a single antenna is employed at the UE to keep the complexity of the front-end limited. The framework described in the paper can be generalized to the case of a multicellular HetNets scenario (including MIMO configurations) in a straightforward manner.

Let $\mathbf{h}_{kj,n} \in \mathbb{C}^{M_{\psi(k)} \times 1}$ denote the uplink channel vector with entries $[\mathbf{h}_{kj,n}]_m$ representing the (frequency) channel gains over subcarrier n from the j-th UE to the m-th receive antenna of the serving AP $\psi(k)$ of user k, where $\psi(k) : \mathcal{K} \mapsto \mathcal{S}$ is a generic function that assigns each user k its serving AP.[1] In the following, $\mathcal{K} = \{1, \ldots, K\}$ and $K = \sum_{s=0}^{S} K_s$ denote the set and the number of UE in the network respectively, with K_s representing the number of UE in the s-th cell: if $s = 0$, the UE will be termed macrocell user equipment (MUE), and small-cell user equipment (SUE) otherwise, although there is no substantial distinction among the two classes of users (this is clarified further in the rest of this paper). We also assume that the channels remain constant within a reasonable time interval (for more quantitative details, see Section V).

We let $z_{j,n}$ denote the data symbol of UE j over subcarrier n and write $p_{j,n}$ for its corresponding power. The vector $\mathbf{x}_{k,n} \in \mathbb{C}^{M_{\psi(k)} \times 1}$ collecting the samples received over subcarrier n at the AP serving the k-th UE can thus be written as

$$\mathbf{x}_{k,n} = \sqrt{p_{k,n}}\mathbf{h}_{kk,n}z_{k,n} + I_{k,n} + \mathbf{w}_{k,n} \tag{1}$$

where $\mathbf{w}_{k,n} \sim \mathcal{CN}(\mathbf{0}_{M_{\psi(k)}}, \sigma^2\mathbf{I}_{M_{\psi(k)}})$ is thermal noise and

$$I_{k,n} = \sum_{j=1, j\neq k}^{K} \sqrt{p_{j,n}}\mathbf{h}_{kj,n}z_{j,n} \tag{2}$$

accounts for the multiple access interference (MAI) experienced by user k over subcarrier n. Note that $I_{k,n}$ accounts for both intra-cell interference (generated by other UEs served by the same AP) and inter-cell interference (from UEs served by all other APs). To keep the complexity at a tolerable level, a simple linear detection scheme is employed for data detection, although a generalization to nonlinear detectors is straightforward. This means that the entries of $\mathbf{x}_{k,n}$ are linearly combined to form $y_{k,n} = \mathbf{g}_{k,n}^H\mathbf{x}_{k,n}$ where $\mathbf{g}_{k,n}$ is the vector employed for recovering the data transmitted by user k over subcarrier n. Then, the signal-to-interference-plus-noise ratio (SINR) over the n-th subcarrier that is achieved by user k at its serving AP takes the form:

$$\gamma_{k,n} = \mu_{k,n}(\mathbf{p}_{-k,n})p_{k,n} \tag{3}$$

[1]For a more detailed description of this assignment mapping, see Section V.

where $\mathbf{p}_{-k,n} = (p_{1,n}, \ldots, p_{k-1,n}, p_{k+1,n}, \ldots, p_{K,n})^T$ denotes the power profile of all users except k over subcarrier n, and

$$\mu_{k,n}(\mathbf{p}_{-k,n}) = \frac{\left|\mathbf{g}_{k,n}^H\mathbf{h}_{kk,n}\right|^2}{\|\mathbf{g}_{k,n}\|^2\sigma^2 + \sum_{j=1,j\neq k}^{K}\left|\mathbf{g}_{k,n}^H\mathbf{h}_{kj,n}\right|^2 p_{j,n}}. \tag{4}$$

Using (3), the achievable rate (normalized to the subcarrier bandwidth, and thus measured in b/s/Hz) of the k-th user will be:

$$r_k(\mathbf{p}) = \frac{1}{N}\sum_{n=1}^{N}\log_2(1 + \gamma_{k,n}) \tag{5}$$

where $\mathbf{p}_k = (p_{k,1}, \ldots, p_{k,N})$ denotes the power profile of user k over all subcarriers $n = 1, \ldots, N$, and $\mathbf{p} = (\mathbf{p}_1, \ldots, \mathbf{p}_K) \in \mathbb{R}_+^{KN}$ is the corresponding power profile of all users (obviously, $p_{k,n} = 0$ if user k is not transmitting over subcarrier n). To simplify notation, the argument of $\mu_{k,n}$ and r_k will be suppressed in what follows.

B. Problem Formulation

As mentioned in Section I, *energy-efficient* network design must take into account the energy consumption incurred by each UE. To that end, note that, in addition to the radiated powers $\mathbf{p}_k$ at the output of the radio-frequency front-end, each terminal k also incurs circuit power consumption during transmission, mostly because of power dissipated at the UE signal amplifier [21], [23], [33]. Therefore, the overall power consumption $P_{T,k}$ of the k-th UE will be given by

$$P_{T,k} = p_{c,k} + P_k = p_{c,k} + \sum_{n=1}^{N} p_{k,n} \tag{6}$$

where $P_k = \sum_{n=1}^{N} p_{k,n}$ is the transmitted power of user k over the entire spectrum, while $p_{c,k}$ represents the average power consumed by the device electronics of the k-th UE (assumed for simplicity to be independent of the transmission state). Following [23], [34], the *energy efficiency* of the link can then be measured (in b/J/Hz) by the utility function

$$u_k(\mathbf{p}) = \frac{r_k}{P_{T,k}} = \frac{N^{-1}\sum_{n=1}^{N}\log_2(1 + \mu_{k,n}p_{k,n})}{p_{c,k} + \sum_{n=1}^{N} p_{k,n}} \tag{7}$$

where the dependence on the transmit power vectors of all other users is subsumed in the gains $\boldsymbol{\mu}_k = \{\mu_{k,n}\}_{n=1}^N$ of (4). Accordingly, in data-oriented wireless networks, QoS requirements take the form $r_k \geq \theta_k$, where θ_k is the minimum rate threshold required by user k.

To summarize, the design of an energy-efficient resource allocation scheme which encompasses both subcarrier allocation and power control amounts to solving the following multi-agent, multi-objective optimization problem:

$$\text{maximize} \quad u_k(\mathbf{p}) \tag{8a}$$

$$\text{subject to} \quad N^{-1}\sum_{n=1}^{N}\log_2(1 + \mu_{k,n}p_{k,n}) \geq \theta_k \tag{8b}$$

where $u_k(\mathbf{p})$ is the energy efficiency utility function (7) and (8b) represents the normalized rate requirement. Thus, unlike other OFDMA resource allocation problems (such as [35], [36]), subcarrier selection and power loading are tackled in

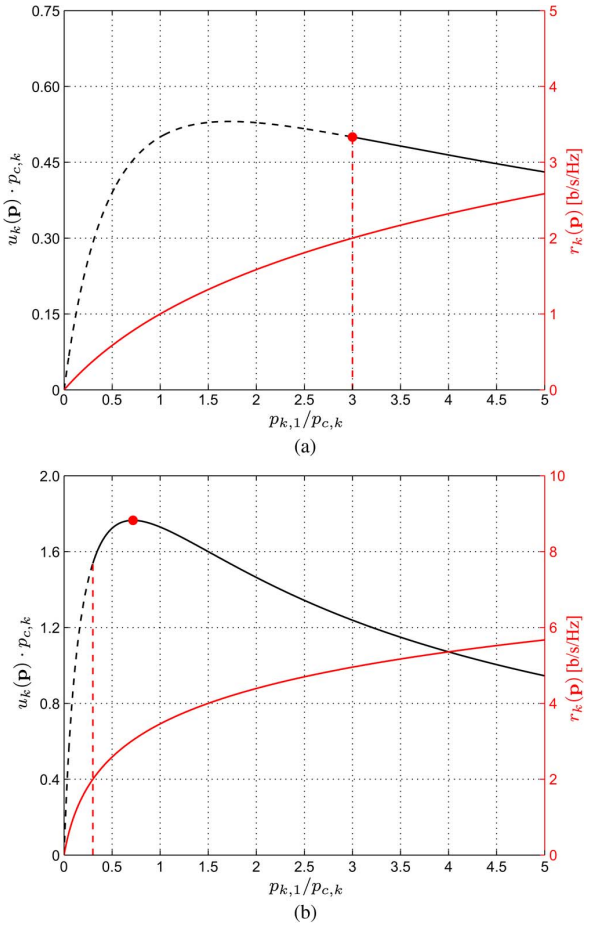


Fig. 1. Normalized utility as a function of the normalized transmit powers ($N = 1$, $\theta_k = 2$ b/s/Hz). (a) $\mu_k \cdot p_{c,k} = 1$. (b) $\mu_k \cdot p_{c,k} = 10$.

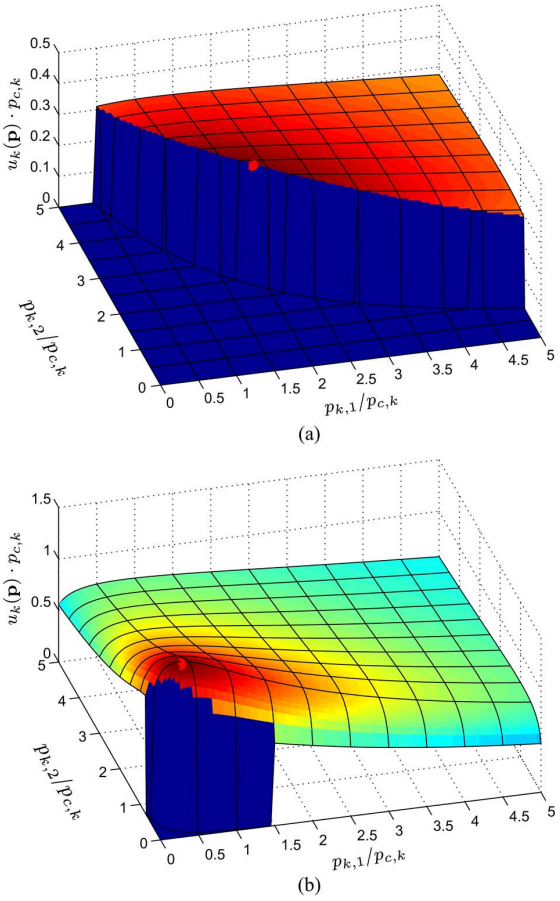


Fig. 2. Normalized utility as a function of the normalized transmit powers ($N = 2$, $\theta_k = 2$ b/s/Hz). (a) $\boldsymbol{\mu}_k \cdot p_{c,k} = (1, 2)$. (b) $\boldsymbol{\mu}_k \cdot p_{c,k} = (10, 20)$.

a *joint* manner. Furthermore, inter- and intra-cell interference between UE transforms (8) into a game where each UE $k \in \mathcal{K}$ aims at unilaterally maximizing its individual link energy-efficiency via an optimal choice of power allocation vector $\mathbf{p}_k$—and, in so doing, obviously affects the possible choices of all other UE in the network.

Remark 1: To visualize the impact of the rate constraints (8b) on the optimization problem (8), Figs. 1 and 2 depict the graph of the utility function (7) of user k (normalized by $p_{c,k}$) as a function of the transmit powers $\mathbf{p}_k = \{p_{k,n}\}_{n=1}^N$ for a fixed interference power vector $\mathbf{p}_{-k}$ (and hence keeping $\{\mu_{k,n}(\mathbf{p}_{-k})\}_{n=1}^N$ fixed). For the sake of visualization, Fig. 1 depicts only $N = 1$ subcarrier. The dashed black line depicts the unconstrained utility (7), whereas the solid black line reports $u_k(\mathbf{p})$ for the values of $p_{k,1}$ such that (8b) holds, assuming $\theta_k = 2$ b/s/Hz (for convenience, also the rate r_k is reported with red lines): $\mu_{k,1} = 1/p_{c,k}$ in Fig. 1(a), whereas $\mu_{k,1} = 10/p_{c,k}$ in Fig. 1(b). As can be seen, the power level that maximizes $u_k(\mathbf{p})$ (red dot) is on the left boundary of the feasible power set of Fig. 1(a): in this case, maximizing $u_k(\mathbf{p})$ corresponds to minimizing the power subject to rate constraints, e.g., as considered in [30]. In general however, the maximization of energy efficiency produces a different optimal point, as reported in Fig. 1(b) where the focal user can exploit better channel conditions experienced to increase its utility. This formulation is particularly appealing for next-generation wireless systems [28], as it captures the tradeoff between obtaining a satisfactory spectral efficiency and saving as much energy as possible [20], [23], [34]. This behavior is analogous to what can be observed in Fig. 2 where $N = 2$ and $\theta_k = 2$ b/s/Hz. When the channel conditions are not favorable (in Fig. 2(a), $\boldsymbol{\mu}_k \cdot p_{c,k} = (1, 2)$), the optimal power allocation $\mathbf{p}_k/p_{c,k} = (1.83, 2.33)$ lies on the contour of the (normalized) utility surface that guarantees $r_k(\mathbf{p}) \geq \theta_k$ (when $r_k(\mathbf{p}) < \theta_k$, we assume here $u_k(\mathbf{p}) = 0$ for the sake of graphical representation)—thus getting $r_k(\mathbf{p}) = \theta_k$. On the contrary, when the channel conditions are more favorable (in Fig. 2(b), $\boldsymbol{\mu}_k \cdot p_{c,k} = (10, 20)$), the utility is maximized by $\mathbf{p}_k/p_{c,k} = (0.37, 0.42)$, that yields $r_k(\mathbf{p}) = 2.74$ b/s/Hz $> \theta_k$.

Remark 2: It is easy to see that a particular set of constraints $\{\theta_k\}_{k=1}^K$ may affect the *feasibility* of the problem in the sense that there might not exist *any* power allocation $\mathbf{p} \in \mathbb{R}_+^{KN}$ that allows all constraints θ_k to be met *simultaneously* — essentially due to mutual interference in the network, which implies a dependence between the gains $\boldsymbol{\mu}_k$ $\forall k$. Necessary and sufficient

conditions that ensure the feasibility of the problem (8) in the single-carrier case $N = 1$ can be found in [22]. On the other hand, analogous conditions for the general case of $N > 1$ subcarriers are very difficult to obtain, and future investigations will focus on addressing this issue.

III. Game-Theoretic Resource Allocation

A. Game-Theoretic Formulation of the Problem

As mentioned earlier, mutual interference in the network introduces interactions among the users that aim at optimizing their utilities (8). A natural framework for studying such strategic inter-user interactions is offered by the theory of non-cooperative games with continuous (and action-dependent) action sets. Thus, following Debreu [25] (see also [26]), we will formulate the problem as a non-cooperative game $\mathcal{G} \equiv \mathcal{G}(\mathcal{K}, \mathcal{P}, u)$ consisting of the following components:

a) The set of *players* of $\mathcal{G}$ is the set $\mathcal{K}$ of the network's UE.
b) A priori, each player can choose any transmit power vector in $\mathcal{P}_k^0 \equiv \mathbb{R}_+^N$. However, given a power profile $\mathbf{p}_{-k} \in \mathcal{P}_k^0 \equiv \prod_{\ell \neq k} \mathcal{P}_\ell^0$ of the opponents of player k, the *feasible action set* of player k in the presence of the rate requirements (8b) is:
$$\mathcal{P}_k(\mathbf{p}_{-k}) = \left\{ \mathbf{p}_k \in \mathcal{P}_k^0 : r_k(\mathbf{p}) \geq \theta_k \right\}. \tag{9}$$
c) The *utility* $u_k(\mathbf{p}_k; \mathbf{p}_{-k})$ of player k is given by (7).

In this framework, the most widely used solution concept is a generalization of the notion of Nash equilibrium [4], known as *Debreu equilibrium* (DE) [25] and sometimes also referred to as *generalized Nash equilibrium* (GNE) [26]. Formally:

Definition 1: A power profile $\mathbf{p}^\star$ is a DE of the energy-efficiency game $\mathcal{G}$ if
$$\mathbf{p}_k^\star \in \mathcal{P}_k(\mathbf{p}_{-k}^\star) \qquad \forall k \in \mathcal{K}, \tag{10a}$$
and
$$u_k(\mathbf{p}^\star) \geq u_k(\mathbf{p}_k; \mathbf{p}_{-k}^\star) \qquad \forall \mathbf{p}_k \in \mathcal{P}_k(\mathbf{p}_{-k}^\star), k \in \mathcal{K}. \tag{10b}$$

The main difference between Debreu and Nash equilibria is that the latter notion posits that players can unilaterally deviate to *any* feasible action, irrespective of whether this action satisfies the (coupled) constraints imposed on a player's action set by the actions of other players in the game. Put differently, Nash-type deviations include any action that satisfies a player's individual, *uncoupled* constraints, even if so doing violates the player's *coupled* constraints. In the case at hand, this means that, at Nash equilibrium, users would be allowed to transmit at any power level, even if this violates the system's transmission rate requirements. On the other hand, these feasibility constraints are already ingrained in the DE concept: the only unilateral deviations considered in (10b) are those for which the rate constraints are satisfied.[2]

[2]The difference between Nash and Debreu equilibria is highlighted further if each player's transmit power is also constrained by a peak value (see below for more details): in this case, each user's individual power constraints would have to be satisfied by Nash-type deviations (and, of course, Debreu-type deviations as well), but Nash-type deviations would not necessarily satisfy the users' coupled QoS constraints.

As such, Debreu equilibria are of particular interest in the context of distributed systems because they offer a stable solution of the game from which players (in this case, UEs) have no incentive to deviate from (and thus destabilize the system) if everyone else maintains their chosen power allocation profiles. Accordingly, in what follows, we investigate the existence and characterization of DE in the energy-efficient power allocation game $\mathcal{G}$, leaving the question of uniqueness and convergence to such states to Sections III-C and IV, respectively.

B. Problem Feasibility and Equilibrium Existence

Debreu's original analysis [25] provides a general equilibrium existence result under the following assumptions:

(D1) The players' feasible action sets $\mathcal{P}_k(\mathbf{p}_{-k})$ are nonempty, closed, convex, and contained in some compact set $\mathcal{C}_k$ for all $\mathbf{p}_{-k} \in \mathcal{P}_{-k} \equiv \prod_{\ell \neq k} \mathcal{P}_\ell$.
(D2) The sets $\mathcal{P}_k(\mathbf{p}_{-k})$ vary continuously with $\mathbf{p}_{-k}$ (in the sense that the graph of the set-valued correspondence $\mathbf{p}_{-k} \mapsto \mathcal{P}_k(\mathbf{p}_{-k})$ is closed).
(D3) Each user's payoff function $u_k(\mathbf{p}_k; \mathbf{p}_{-k})$ is quasi-concave in $\mathbf{p}_k$ for all $\mathbf{p}_{-k} \in \mathcal{P}_{-k}$.

In our setting, $r_k(\mathbf{p}_k; \mathbf{p}_{-k})$ in (5) is concave in $\mathbf{p}_k$ and unbounded from above, so $\mathcal{P}_k(\mathbf{p}_{-k})$ is convex and nonempty for all $\mathbf{p}_{-k} \in \mathcal{P}_k^0$. Moreover, $\mathcal{P}_k(\mathbf{p}_{-k})$ varies continuously with $\mathbf{p}_{-k}$ because the constraints (8b) are themselves continuous in $\mathbf{p}_{-k}$. Finally, it is easy to show that $u_k(\mathbf{p}_k; \mathbf{p}_{-k})$ is quasi-concave in $\mathbf{p}_k$: since $u_k(\mathbf{p}_k; \mathbf{p}_{-k}) \geq a$ if and only if
$$r_k(\mathbf{p}_k; \mathbf{p}_{-k}) - a\left(p_c + \sum\nolimits_{n=1}^{N} p_{k,n}\right) \geq 0 \tag{11}$$
and the set defined by this inequality is convex for every $\mathbf{p}_{-k} \in \mathcal{P}_{-k}$ (recall that r_k is concave in $\mathbf{p}_k$), quasi-concavity of $u_k(\cdot, \mathbf{p}_{-k})$ follows.

However, even though the users' best response sets
$$\mathcal{P}_k^\star(\mathbf{p}_{-k}) \equiv \underset{\mathbf{p}_k \in \mathcal{P}_k(\mathbf{p}_{-k})}{\arg\max}\ u_k(\mathbf{p}_k; \mathbf{p}_{-k}) \tag{12}$$
are nonempty, convex, closed and bounded for every $\mathbf{p}_{-k}$, they might (and typically do) run off to infinity—i.e., they are not *uniformly* bounded. To understand this, simply consider the case of two UEs transmitting over a single channel: if one of the UE transmits at very high power, the other UE is forced to transmit at a commensurately high power in order to meet its rate requirement. This leads to a cascade of power increases that makes each UE's feasible action set $\mathcal{P}_k(\mathbf{p}_{-k})$ (and, hence, $\mathcal{P}_k^\star(\mathbf{p}_{-k})$ as well) escape to infinity as the other UE increases its individual power. Formally, this means that the UE's feasible action sets $\mathcal{P}_k(\mathbf{p}_{-k})$ are not contained in an enveloping bounded set $\mathcal{C}_k$. Thus, Debreu's equilibrium existence theorem [25] does not apply.

From a power control perspective, this is not surprising: as is well known [37], the problem (8) may fail to be feasible, i.e., there may be no power profile $\mathbf{p} = (\mathbf{p}_1, \ldots, \mathbf{p}_K)$ such that $\mathbf{p}_k \in \mathcal{P}_k(\mathbf{p}_{-k})$ for all k. Obviously, in this case, the energy-efficiency game $\mathcal{G}$ does not admit an equilibrium either. On the other hand, at a purely formal level, equilibrium existence and problem feasibility are restored if we assume that users can transmit

with infinitely high power, i.e., each UE $k \in \mathcal{K}$ chooses its total transmit power from the compactified half-line $[0, +\infty]$. In this extended setup, there are two points where indeterminacies may arise: first, the utility of player k is not well-defined if $p_{k,n} = +\infty$ for some n; second, the rate requirement (8b) of user k is also ill-defined if $p_{\ell,n} = +\infty$ for some $\ell \neq k$. To address these problems, note first that the utility function (7) of player k decreases to 0 when $p_{k,n} \to +\infty$ for some channel $n = 1, \ldots, N$, reflecting the fact that $\lim_{x\to+\infty} x^{-1} \log_2 x = 0$. Thus, by continuity, the utility of player k for infinite transmit powers $p_{k,n}$ may be defined as:

$$u_k(\mathbf{p}) = 0 \quad \text{whenever } p_{k,n} = +\infty \text{ for some } n. \tag{13}$$

As for the rate requirements of user k, a simple exponentiation of (8b) for finite p yields the equivalent expression:

$$\prod_{n=1}^{N} (1 + \mu_{k,n} p_{k,n}) \geq 2^{N\theta_k} \tag{14}$$

or, after substituting for $\mu_{k,n}$ and rearranging:

$$\prod_{n=1}^{N} \left(\|\mathbf{g}_{k,n}\|^2 \sigma^2 + \sum_{j=1}^{K} |\mathbf{g}_{k,n} \mathbf{h}_{kj,n}|^2 p_{j,n} \right) \geq 2^{N\theta_k} \prod_{n=1}^{N} \left(\|\mathbf{g}_{k,n}\|^2 \sigma^2 + \sum_{j\neq k} \left|\mathbf{g}_{k,n} \mathbf{h}_{kj,n}\right|^2 p_{j,n} \right). \tag{15}$$

Since both sides of (15) are well-defined for all $p_{j,n} \in [0, +\infty]$, (15) provides a reformulation of (8b) that remains meaningful even in the extended arithmetic of $[0, +\infty]$.

In this infinite-power framework, any power profile $\mathbf{p}^\star = (\mathbf{p}_1^\star, \ldots, \mathbf{p}_K^\star)$ with $\sum_{n=1}^{N} p_{k,n}^\star = +\infty$ for all $k \in \mathcal{K}$ is feasible with respect to (15). Furthermore, if player k deviates *unilaterally* and starts transmitting with finite total power, its rate requirement (15) will be automatically violated and its utility equals 0. Consequently, no player can gain a utility greater than 0 by deviating from $\mathbf{p}^\star$. This shows that the resulting infinite-power game $\overline{\mathcal{G}}$ with utility functions and rate requirements extended as in (13) and (15) above always admits a DE—and trivially so. However, any such equilibrium is clearly unreasonable from a practical standpoint as it represents a cascade of power increases that escapes to infinity as players try to meet their power constraints.

In view of the above, we could consider an alternative formulation of $\mathcal{G}$ in which the users' *uncoupled* action sets (i.e., unadjusted for the actions of other users) are of the form

$$\mathcal{P}_k^0 = \left\{ \mathbf{p}_k \in \mathbb{R}_+^N : 0 \leq p_{k,n} \leq \overline{p}_{k,n}, \sum\nolimits_n p_{k,n} \leq \overline{P}_k \right\} \tag{16}$$

for given maximum per-subcarrier transmit power levels $\overline{p}_{k,n}$ and total power constraints $\overline{P}_k$. In this case however, a crucial arising problem is that the resulting system could be even *unilaterally infeasible* in the sense that the admissible action set $\mathcal{P}_k(\mathbf{p}_{-k})$ of player k may be empty for a wide range of transmit power profiles $\mathbf{p}_{-k}$ of the other users in the system. Put differently, in the presence of maximum power constraints (a case that will be discussed at the end of Section IV), any given user may not be able to even participate in the game (in contrast with the formulation (9) of $\mathcal{G}$), thus exacerbating the equilibrium existence problem.

Of course, given that actual wireless devices cannot transmit at arbitrarily high levels, it is still crucial to determine under which conditions the game $\mathcal{G}$ admits a realizable DE. Therefore, in what follows, we will focus on conditions and scenarios, which guarantee that:

1) The energy-efficiency game $\mathcal{G}$ admits a DE with *finite* transmit powers (Section III-C).
2) This equilibrium is unique (Section III-C).
3) Users converge to equilibrium by following an adaptive, distributed algorithm (Section IV).

C. Equilibrium Characterization and Uniqueness

The goal of this section is to characterize the game's DE by exploiting the fact that they are the fixed points of a certain best-response mapping.

Proposition 1: A transmit power profile $\mathbf{p}^\star$ is a DE if and only if its components $p_{k,n}^\star$ satisfy:

$$p_{k,n}^\star = \left(\frac{1}{\lambda_k^\star} - \frac{1}{\mu_{k,n}} \right)^+ \tag{17}$$

where

$$\lambda_k^\star = \min\left\{ \lambda_k, \overline{\lambda}_k \right\}. \tag{18}$$

In the above,

$$\lambda_k = \frac{W\left(\alpha_k \cdot e^{\beta_k - 1}\right)}{\alpha_k} \tag{19}$$

is the water level of the WF operator (17) when the problem (8) is solved without the minimum-rate constraints (8b) (i.e., when $\theta_k = 0$ for all $k \in \mathcal{K}$), $W(\cdot)$ denotes the Lambert W function [32], while

$$\alpha_k = |\mathcal{S}_k|^{-1} \left(p_{c,k} - \sum_{n\in\mathcal{S}_k} \mu_{k,n}^{-1} \right) \tag{20}$$

and

$$\beta_k = \mathcal{S}_k^{-1} \sum_{n\in\mathcal{S}_k} \ln \mu_{k,n} \tag{21}$$

where $\mathcal{S}_k = \left\{ n \in \mathcal{N} : \mu_{k,n} \geq \lambda_k \right\}$ denotes the subset of active subcarriers when using the uncostrained energy-efficient formulation. Similarly:

$$\overline{\lambda}_k = \left(2^{-N\theta_k} \prod\nolimits_{n\in\overline{\mathcal{S}}_k} \mu_{k,n} \right)^{1/|\overline{\mathcal{S}}_k|} \tag{22}$$

is the water level of (17) when all minimum-rate constraints (8b) are met simultaneously with equality (i.e., (8) reduces to a power minimization problem with equality rate constraints $r_k = \theta_k$), and, as above, $\overline{\mathcal{S}}_k = \left\{ n \in \mathcal{N} : \mu_{k,n} \geq \overline{\lambda}_k \right\}$ denotes the subset of active subcarriers.

Proof: The proof is given in Appendix A and relies on defining the best-response mapping and using fractional programming to characterize its fixed points. ■

Remark 3: Proposition 1 does not provide a way to calculate the water levels λ_k and $\overline{\lambda}_k$. For an iterative computational method, the reader is referred to Section IV.

Despite its convoluted appearance, Proposition 1 is of critical importance from both a theoretical and practical point of view. Indeed, it is the basic step to derive sufficient conditions ensuring the existence and uniqueness of the DE and also to develop a distributed and scalable power allocation algorithm that steers the network to a stable equilibrium state.

To that end, note that the equilibrium characterization of Proposition 1 may be vacuous if the game does not admit a DE to begin with—for instance, if the original power control problem is not feasible. On that account, we have:

Proposition 2: The energy-efficiency game $\mathcal{G}$ admits a unique DE $\mathbf{p}^\star$ whenever $\forall k \in \mathcal{K}$:

$$\sum_{\substack{j=1\\ j\neq k}}^{K}\sum_{n=1}^{N}\omega_{kj,n}^2 \sup_{\boldsymbol{\mu}_k\in\boldsymbol{\Omega}_k}\left[\frac{1}{\varsigma_k^\star}\sum_{n\in\mathcal{S}_k^\star}\omega_{kk,n}^{-2}\left(\xi_{k,n}^2+\varsigma_k^\star-2\xi_{k,n}\right)\right]<1 \tag{23}$$

where $\boldsymbol{\Omega}_k=\prod_{n=1}^{N}(0,\sigma^{-2}\omega_{kk,n}]$, $\varsigma_k^\star=|\mathcal{S}_k^\star|$,

$$\omega_{kj,n}=\frac{|\mathbf{g}_{k,n}^H\mathbf{h}_{kj,n}|^2}{||\mathbf{g}_{k,n}||^2} \tag{24}$$

and

$$\mathcal{S}_k^\star=\begin{cases}\overline{\mathcal{S}}_k & if \lambda_k\geq\overline{\lambda}_k\\ \mathcal{S}_k & if \lambda_k<\overline{\lambda}_k\end{cases} \tag{25}$$

$$\xi_{k,n}=\begin{cases}\mu_{k,n}\overline{\lambda}_k^{-1} & \text{if } \overline{\lambda}_k\leq\lambda_k \text{ and } n\in\mathcal{S}_k^\star\\ \frac{\mu_{k,n}-\lambda_k}{\lambda_k(1+\nu_k)} & \text{if } \overline{\lambda}_k>\lambda_k \text{ and } n\in\mathcal{S}_k^\star\\ 0 & \text{if } n\notin\mathcal{S}_k^\star\end{cases} \tag{26}$$

with $\nu_k=-\ln\lambda_k+(\beta_k-1)$.

Proof: The main steps for the proof are given in Appendices B and C; for a more detailed version, the reader is referred to the online technical report [38]. ■

Remark 4: Notice that these sufficient conditions are similar to the well-known conditions ensuring the uniqueness of a Nash equilibrium in the non-cooperative rate maximization game studied by [9] in the context of the interference channel. Intuitively, (23) means that if the interfering connections for a user are sufficiently far away and the resulting SINR is high enough, then the DE exists and is unique. However, these conditions include a non-trivial optimization step w.r.t. $\boldsymbol{\mu}_k$ that depends on the actual opponents' power $\mathbf{p}_{-k}$. Indeed, the variables of the problem impact the values of $\lambda_k^\star$, $\mathcal{S}_k^\star$ and all functions $\xi_{k,n}$, making the conditions rather difficult to be exploited. To tackle this issue, the online technical report [38] provides a set of sufficient conditions that are simpler. This is achieved by observing that the upper-bound of the supremum term in (23) boils down to computing a function of the system parameters only. The downside is that these simple conditions are more stringent than (23). Nevertheless, it is worth pointing out that the users of the network are never required to compute these conditions: (23) only meant as a safety feature to guard against catastrophic system instabilities, to be calculated by the network administrator based on expected network usage scenarios.

Remark 5: Since the conditions of Proposition 2 are only sufficient, DE might exist even in the case where (23) does not hold for some $k\in\mathcal{K}$. As a matter of fact, when (8) is feasible, the distributed algorithm that we present in Section IV is observed to converge to a DE in all the numerical simulations performed and for every network scenario considered.

IV. Distributed Implementation

To derive a practical procedure allowing UE to reach the DE of $\mathcal{G}$ in a distributed fashion (*without* any distinction between SUE and MUE), we start by focusing on a specific UE $k\in\mathcal{K}$ and assume that all other UE $j\neq k$ have already chosen their optimal transmit powers $\mathbf{p}_{-k}=\mathbf{p}_{-k}^\star$ (in a possibly asynchronous fashion). From (4), we then see that the gains $\mu_{k,n}(\mathbf{p}_{-k,n}^\star)$ needed to implement (17) are simply

$$\mu_{k,n}(\mathbf{p}_{-k,n}^\star)=\frac{\gamma_{k,n}}{p_{k,n}} \tag{27}$$

for all $n\in\mathcal{N}$. This means that the only information that is not locally available at the k-th UE to compute the optimal powers $\{p_{k,n}^\star\}$ is the set of SINR $\{\gamma_{k,n}\}$ measured at the serving SCA of the k-th UE, and which can be sent with a modest feedback rate requirement on the return channel (a discussion on the impact of a limited feedback can be adapted to this specific scenario from [39]).

Based on the above considerations, we can derive an iterative and fully decentralized algorithm to be adopted by each UE k at each time step t to solve the fixed-point system of equations (17) with a low-complexity, scalable and adaptive procedure. The pseudocode for the whole network is summarized in Algorithm 1. Note that, in practice, each UE $k\in\mathcal{K}$ only needs to implement the steps for only one value in the user loop (i.e., its own index), so the algorithm is suitable for asynchronous implementation in dynamic network configurations where each UE only requires the SINR to be fed back by the serving SCA, without any further information on the network.

Algorithm 1 Iterative Algorithm to Solve Problem (8).

set $t=0$
initialize $\mathbf{p}_k[t]=\mathbf{0}_N$ for all users $k\in\mathcal{K}$
repeat
  for $k=1$ to K **do**
    {loop over the users}
    receive $\{\gamma_{k,n}[t]\}_{n=1}^N$ from the serving AP
    compute λ_k using Algorithm 2 and $\overline{\lambda}_k$ using inverse water-filling
    set $\lambda_k^\star=\min\{\lambda_k,\overline{\lambda}_k\}$
    for $n=1$ to N **do**
      {loop over the carriers}
      update $p_{k,n}[t+1]=(1/\lambda_k^\star-p_{k,n}[t]/\gamma_{k,n}[t])^+$
    end for
  end for
  update $t=t+1$
until $\mathbf{p}_k[t]=\mathbf{p}_k[t-1]$ for all $k\in\mathcal{K}$

For the sake of clarity, the algorithm to compute λ_k for each UE $k \in \mathcal{K}$ as in (19) is reported in Algorithm 2, whereas $\overline{\lambda}_k$ can easily be computed using standard iterative water-filling (IWF) methods [27]. Note that, although (19) is derived analytically in closed form and can be computed directly, it is still appealing to use the iterative procedure outlined in Algorithm 2, which takes advantage of the Dinkelbach approach [40] based on Newton's method. The latter is known to converge superlinearly for convex nonlinear fractional programming problems [40], and leads to substantial computational savings compared to evaluating the Lambert function directly. Interestingly, the Dinkelbach algorithm can also be properly modified to address the computation of the IWF-based quantity $\overline{\lambda}_k$, thus saving the complexity required for sorting the coefficients $\{\mu_{k,n}\}_{n=1}^{N}$ in a descending order [41]. For the sake of brevity, Algorithm 2 makes use of some functions that are introduced in the proof of Proposition 1 (Appendix A). For future reference, throughout the simulations reported in Section V, the convergence tolerance is set to $\varepsilon = 10^{-5}$, and we check whether the end state of the algorithm is a DE by testing the characterization of Proposition 1.

Algorithm 2 Iterative Algorithm to Compute λ_k as in (19).

set a tolerance $\varepsilon \ll 1$
{initialization of the Dinkelbach method:}
repeat
  select a random $\lambda_k \in \mathbb{R}$
  for $n = 1$ to N **do**
    set $p_{k,n} = (1/\lambda_k - p_{k,n}[t]/\gamma_{k,n}[t])^+$
  end for
  compute $\varphi(\mathbf{p}_k)$ and $\chi(\mathbf{p}_k)$ using (31) (see Appendix A)
  set $\Phi(\lambda_k) = \varphi(\mathbf{p}_k) - \lambda_k \chi(\mathbf{p}_k)$
until $\Phi(\lambda_k) \geq 0$
{Dinkelbach method:}
while $\Phi(\lambda_k) \geq \varepsilon$ **do**
  set $\lambda_k = \varphi(\mathbf{p}_k)/\chi(\mathbf{p}_k)$
  for $n = 1$ to N **do**
    set $p_{k,n} = (1/\lambda_k - p_{k,n}[t]/\gamma_{k,n}[t])^+$
  end for
  update $\varphi(\mathbf{p}_k)$ and $\chi(\mathbf{p}_k)$ using (31)
  set $\Phi(\lambda_k) = \varphi(\mathbf{p}_k) - \lambda_k \chi(\mathbf{p}_k)$
end while

Proposition 3: The iterates of Algorithm 1 converge to DE whenever (23) holds.

Proof: The convergence of Algorithm 1 to an equilibrium point follows from the contraction properties of the best-response mapping investigated in Section III-C. ■

Remark 6: Although the contraction properties of the best-response mapping are contingent on the sufficient conditions of Proposition 2, Algorithm 1 is still seen to converge to a DE of $\mathcal{G}$, provided that the problem is feasible to begin with (see the next section for a numerical assessment via extensive numerical simulations).

Remark 7: In the theoretical analysis of Section III (as well as in Algorithm 1), we consider neither total maximum power constraints $\overline{P}_k$, such that, $P_k \leq \overline{P}_k$, nor per-subcarrier maximum power constraints $\overline{p}_{k,n}$, such that $p_{k,n} \leq \overline{p}_{k,n}$. Although power masks are usually required by wireless standards to meet out-of-band emission policies, the power limits $\{\overline{P}_k\}_{k\in\mathcal{K}}$ and $\{\overline{p}_{k,n}\}_{k\in\mathcal{K},n\in\mathcal{N}}$ significantly impact the analytical characterization of the DE $\mathbf{p}^\star$. For the sake of theoretical correctness, they are thus not included in the present work and are left as a future direction of research. However, it is worth stressing that: i) Algorithm 1 can easily accommodate $\{\overline{P}_k\}_{k\in\mathcal{K}}$ and $\{\overline{p}_{k,n}\}_{k\in\mathcal{K},n\in\mathcal{N}}$, by setting $\lambda_k^\star = \max\{\min\{\lambda_k, \overline{\lambda}_k\}, \underline{\lambda}_k\}$, where $\underline{\lambda}_k$ is computed using direct WF [27] (by maximizing the rate $r_k(\mathbf{p})$ under the constraint $\sum_{n=1}^{N} p_{k,n} = \overline{P}_k$), and by setting

$$p_{k,n}[t+1] = \min\left\{\overline{p}_{k,n}, \left(1/\lambda_k^\star - p_{k,n}[t]/\gamma_{k,n}[t]\right)^+\right\}; \quad (28)$$

ii) reasonable values of $\{\overline{P}_k\}_{k\in\mathcal{K}}$ and $\{\overline{p}_{k,n}\}_{k\in\mathcal{K},n\in\mathcal{N}}$ do not modify the optimal power allocation $\mathbf{p}^\star$ in practice. In the interest of providing a practical algorithm that can be used in real-world scenarios, our extensive simulations in Section V make use of the modified algorithm, in which we observe that the selected values for the power constraints are never active in practice, so the theoretical results of Section III remain valid.

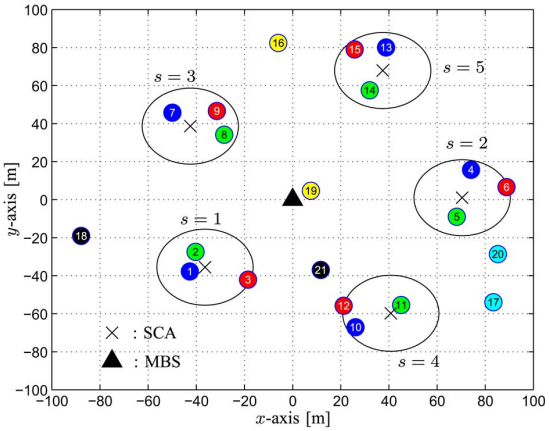


Fig. 3. Random realization of a network with $S = 5$ small cells, $K_S = 3$ SUE, and $K_0 = 6$ MUE, sharing $N = 12$ subcarriers.

V. Numerical Results

Numerical simulations are now used to assess the performance of the proposed algorithm under different operating conditions. To keep the complexity of the simulations tractable while considering a significantly loaded system, we focus on the scenario reported in Fig. 3, where a square-shaped macrocell with an area of 200×200 m^2 centered around its MBS accommodates S randomly distributed small cells, each with a radius of $\rho_s = \rho_S = 20$ m. Throughout the simulations, unless otherwise specified, we adopt the parameters reported in Table I (see [21] and references therein), where, for simplicity, each SC is assumed to have the same number of antennas M_S and to serve the same number of users K_S. Moreover, all UE are assumed to have the same non-radiative power consumption $p_{c,k} = p_c$, and the same power limits $\overline{P}_k = \overline{P}$ and $\overline{p}_{k,n} = \overline{p}$ are imposed for all subcarriers and for all users (see Remark 7).

TABLE I
GENERAL SYSTEM PARAMETERS

Parameter	Value	Parameter	Value
Bandwidth	$B = 11.2$ MHz	Carrier spacing	$\Delta f = 10.9375$ kHz
Carrier frequency	$f_c = 2.4$ GHz	Macro-cell area	0.04 km^2
Total number of small cells	$S = 5$	Small-cell radius	$\rho_S = 20$ m
Number of antennas (MBS, SCA)	$M_0 = 16$, $M_S = 4$	Density of population	1,000 users/km^2
Number of SUE per small cell	$K_S = 4$	Number of MUE	$K_0 = 20$
Number of subcarriers	$N = 96$	Noise power	$B\sigma^2 = -103.3$ dBm
Non-radiative power	$p_c = 20$ dBm	Path-loss exponent	$\zeta = 3.5$
Cut-off parameter	$d_{\text{ref}} = 35$ m	Average path-loss attenuation at d_{ref}	$L_{\text{ref}} = -84.0$ dB
Maximum total power	$\overline{P} = 40$ dBm	Maximum per-subcarrier power	$\overline{p} = 30$ dBm

To include the effects of fading and shadowing, we use the path-loss model introduced in [42], using a 24-tap channel model to reproduce multipath effects. We also assume perfect channel estimation at the receiver end and the use of maximum ratio combining (MRC) techniques, which amounts to setting $\mathbf{g}_{k,n} = \mathbf{h}_{kk,n}$ for all $k \in \mathcal{K}$ and $n \in \mathcal{N}$. The UE $k \in \mathcal{K}$ is then assigned to APs $s \in \mathcal{S}$ following the mapping:

$$\psi(k) = \begin{cases} s & \exists\, s > 0 \text{ s.t. } d_{k,s} \leq \rho_S \\ 0 & \text{otherwise} \end{cases} \tag{29}$$

where $d_{k,s}$ denotes the distance between UE k and SCA s. Without loss of generality, we measure the performance for a specific user (say user 1) within either an SC or a macro-cell, by averaging over all possible positions of the users, uniformly randomizing their minimum-rate constraints θ_k in [0, 2] b/s/Hz for $k \neq 1$.

To evaluate the proposed algorithm in a practical setting, Fig. 3 reports a random realization of the network with the parameters described above, in which the following quantities have been reduced for the sake of graphical representation: $K_S = 3$, $K_0 = 6$, and $N = 12$, $\theta_k = 1.5$ b/s/Hz for SUE, and $\theta_k = 0.5$ b/s/Hz for the MUE. Using the distributed algorithm described in Section IV, after roughly 20 iterations we get the solution to (8), representing the users' power profile at the DE of $\mathcal{G}$, and reported in Fig. 4. Here, the first five subplots correspond to the powers allocated in the small cells (the s-th subplot depicts the powers allocated by the users in the s-th small cell, with colors matching the ones used in Fig. 3), whereas the last two subplots show the powers selected by the MUE labeled 16, 17, 18 (in the sixth subplot) and 19, 20, 21 (in the seventh subplot), respectively. As can be seen in Fig. 4, this method tends to allocate the subcarriers in an exclusive manner whenever the MAI across UE within the same small cell is too large (e.g., see the 4th small cell, in which only 5 subcarriers are shared by the 3 users), and to share the same subcarrier when the MAI across users is at a tolerable level (which also includes the interference generated by SUE from neighboring cells and the MUE). On the right hand side, we report the achieved rates at the DE in b/s/Hz. As can be verified, all users achieve their minimum demands, while for users with particularly favorable channel conditions (in this case, users no. 1, 11, 19, and 21), it is convenient to increase their transmit power so as to obtain better performance in terms of EE. As we mentioned in Section II, we assume the channel

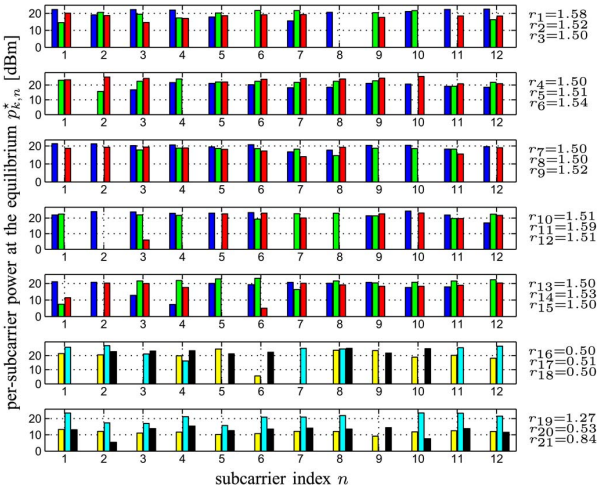


Fig. 4. Outcome of the resource allocation for the scenario of Fig. 3. The subcarriers are allocated exclusively when the MAI within the small cell is large. All users achieve their rate requirements. Users with favorable channels increase their powers to maximize their own utilities.

to be weakly time-varying. Otherwise stated, we assume that the convergence of the proposed algorithm is achieved before significant channel variations, as is customarily assumed in all closed-loop resource allocation schemes. To support this, assume that the uplink and downlink slot durations are in the order of few milliseconds (which is reasonable for LTE/LTE-A standards [43]). In these circumstances, the average convergence time of the proposed solution turns out to be in the order of tens of milliseconds (since convergence is achieved after approximately 20 iterations): such interval is sufficiently shorter than typical channel coherence times, especially when considering usual SC scenarios with pedestrian users.

To assess the robustness of the proposed solution to network perturbations, we depict in Fig. 5 the total power consumption as a function of the iteration step for the network setting of Fig. 3 (lines are identified by UE labels, using the numbering adopted in Fig. 3). In particular, for the sake of clarity, since all other users show similar results, we only report the behavior of SUE in small cells $s = 1$ and $s = 4$, and the MUE 19 and 21, when, at $t = 25$, two cell-edge users (namely, users 3, 12) simultaneously change their receiver association: both become served by the MBS, due to a variation in the received signal strength (with ensuing reduction of their data rate requirements

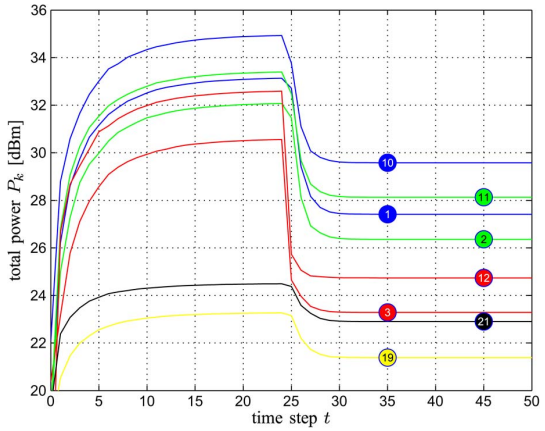


Fig. 5. UE total power consumption as a function of the iteration step. The power allocation fastly converge even in the presence of sudden changes in the network configuration, e.g., due to UE mobility or channel fluctuations.

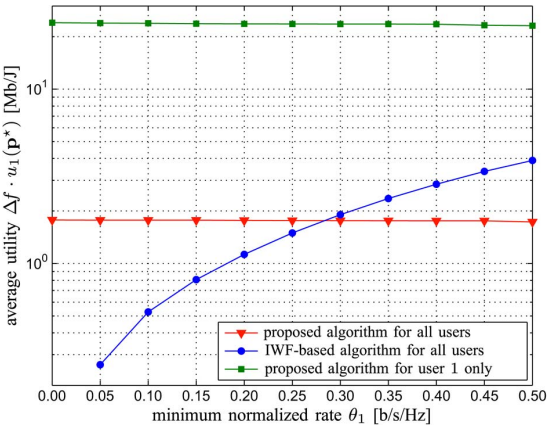


Fig. 6. Average utility at the equilibrium as a function of the minimum rate θ_1. Compared to an IWF-based solution, the Debreu equilibrium may perform worse in terms of overall network utility. However, the IWF-based solution is not a stable operating point: user 1 has always an incentive to deviate and highly increase its own utility.

to 0.5 b/s/Hz, like all other MUE). As can be seen, the algorithm is very robust to network perturbations, and guarantees fast convergence for all users in the network to the new equilibrium point. In this particular example, each UE's power decrease is due to a lower interference generated by the "new" MUE—which, in turn, is a consequence of their lower target rates.

To the best of our knowledge, there are no resource allocation algorithms that address the energy-efficient formulation (8) subject to the minimum-rate demands (8b). To evaluate the improvement in terms of EE of the proposed technique (red), we thus compare its performance with that achieved by an IWF-based solution (blue), in which all users aim at meeting θ_k with equality [30]. Fig. 6 reports the average utility achieved by averaging over all possible positions of a particular MUE (say user 1) as a function of a specific minimum rate θ_1, using

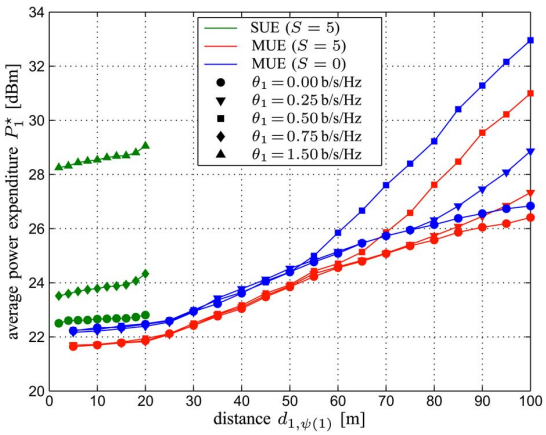


Fig. 7. Average transmit power at the equilibrium as a function of the distance from the receiver. The HetNet configuration ($S = 5$) significantly reduces the power consumption of the UE compared to the macro-cell classical scenario ($S = 0$) for any rate requirements.

the parameters reported in Table I.[3] Interestingly, there exists a critical θ_1 (in this case, 0.28 b/s/Hz), for which the EE of IWF is higher than that achieved by the proposed formulation, mainly due to a weaker MAI caused by the IWF users, that transmit at lower powers than energy-efficient ones (not reported for the sake of brevity). However, IWF policies are *not* stable: if the network's UE adopt an IWF approach, then a UE that deviates from this criterion would *greatly increase its EE* (represented by the green line in Fig. 6). This situation is reminiscent of the well-known prisoner's dilemma [4] where there exist states with higher average utility, but which are obviously abandoned once a user deviates in order to maximize his individual benefits—and, hence, are inherently unstable in a non-cooperative, decentralized setting. In addition to this, the proposed approach shows two interesting properties compared to IWF: *i*) averaging over all network realizations and all minimum rates, Algorithm 1 achieves an average utility of 1.76 Mb/J, which is larger than the IWF-based one, equal to 1.69 Mb/J; and *ii*) it introduces fairness among the users, as its performance in terms of EE is weakly dependent on the QoS requirement θ_k.

To measure the benefits of a HetNet configuration with respect to a classical macrocellular architecture ($S = 0$), Figs. 7 and 8 depict the average total transmit powers and the achievable rates at equilibrium in terms of the distance between the observed user and its receiver, averaged over 2000 independent feasible network realizations per marker. The green and red lines represent the performance in the case of $S = 5$ small cells, $K_S = 4$ SUE, and $K_0 = 20$ MUE, achieved by an SUE and an MUE, respectively, whereas blue lines show the performance obtained by an MUE in the case $S = 0$. We consider three different minimum demands for the SUE (0, 0.75, and 1.5 b/s/Hz,

[3]Throughout all the simulations in the present and subsequent graphs, the selected parameters yield an occurrence of feasible scenarios, assessed a posteriori by letting each UE achieve their minimum-rate constraint (8b) with equality, larger than 99%. Once the scenario is checked to be feasible, the convergence of Algorithm 1 to a stationary point (a DE) occurs with probability 1.

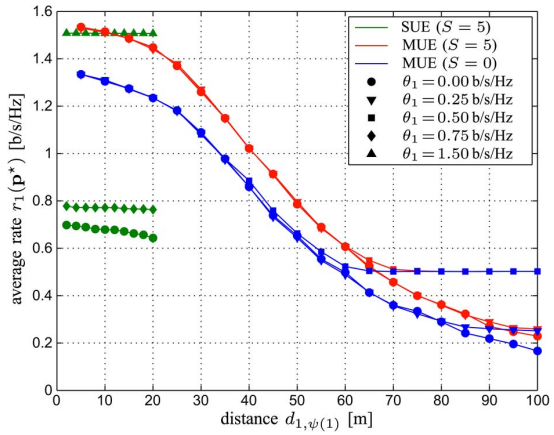


Fig. 8. Average rate at the equilibrium as a function of the distance from the receiver. The HN configuration ($S=5$) significantly increases the rates of the UE compared to the macro-cell classical scenario ($S=0$) for any rate requirements.

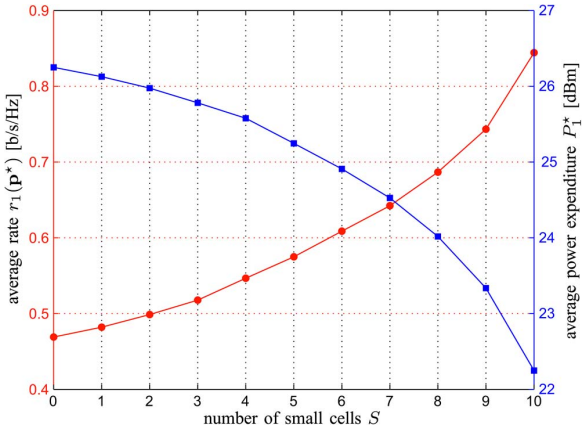


Fig. 9. Average rate at the equilibrium (left axis) and average power consumption (right axis) as functions of the number of small cells. Introducing more small cells increases the average rate and reduces the average power consumption in the network while guaranteeing the minimum rate requirements.

represented by circular, square, and upward-pointing arrowheads), and three different demands for the MUE (0, 0.25, and 0.5 b/s/Hz, represented by circular, downward-pointing arrowheads, and diamond markers respectively). As can be seen, the HetNet configuration introduces *significant gains in both the achievable rates and the power consumption* compared to the classical scenario: by averaging over all possible positions of SUE and MUE across the macrocell area, the MUE get $r_1(\mathbf{p}^\star) \approxeq 0.68$ b/s/Hz with a power consumption $P_1^\star \approxeq 27.5$ dBm (566 mW) when placing $\theta_1 = 0.5$ b/s/Hz,[4] compared to $r_1(\mathbf{p}^\star) \approxeq 0.63$ b/s/Hz with $P_1^\star \approxeq 29.1$ dBm (813 mW) for the same minimum demand in the case $S=0$. The HetNet configuration is also *beneficial in terms of ASE*: using these parameters, we get on average slightly more than 600 b/s/Hz/km^2, compared to 500 b/s/Hz/km^2 for $S=0$.

Introducing small cells has a negative impact in terms of the algorithm's convergence rate: here, on average 4.1 iterations are required for the case $S=5$, compared to 3.5 for the case $S=0$. This is due to decentralizing the resource allocation at each receiving station, thus slightly slowing the convergence of the algorithm. However, this provides a better MAI management ensured by SCAs, that allow SUE to obtain *higher rates with lower interfering powers* at the MBS. As can be seen, due to the path-loss model employed, which is roughly constant for distances within $d_{\text{ref}} > \rho_S$, the SUE performance is *independent of the distance from the SCA*. When SUE place $\theta_1 = 1.5$ b/s/Hz, the spectral efficiency is similar to that achieved by MUE located at comparable distance from the MBS (see Fig. 8), but at the cost of a larger power consumption (see Fig. 7): this is due to a better diversity at the receiver obtained by the MUE, since the MBS employes a larger number of antennas (16 versus 4). However, this does not hold true as the MUE distance increases: averaging over all positions, SUE obtain an average rate $r_1(\mathbf{p}^\star) \approxeq 1.51$ b/s/Hz (more than twice the MUE's one) using $P_1^\star \approx 28.6$ dBm (732 mW, slightly higher than MUE's one).

[4]Note that such minimum demand is about one order of magnitude larger than the one considered for cell-edge users in 4G networks, equal to 0.07 b/s/Hz [43] for a scarcely populated cell (at most 10 users).

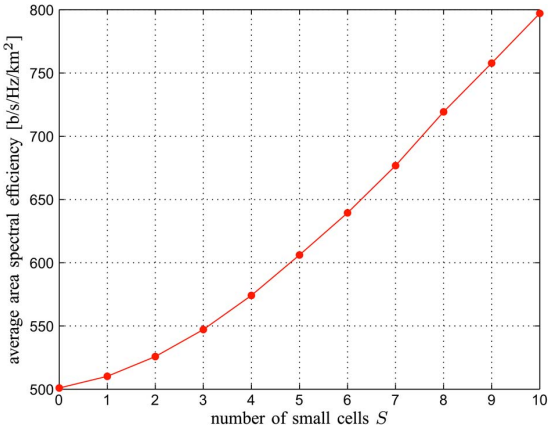


Fig. 10. Average area spectral efficiency as a function of the number of small cells. Introducing more small cells increases the average area spectral efficiency as well.

To emphasize the impact of small cells on the system performance, Figs. 9 and 10 compare the performance, averaged over 10^5 independent network realizations, achieved by an MUE using $\theta_1 = 0.25$ b/s/Hz in the same network as before, populated by $K=40$ users, as a function of the number of SCs S, each having $K_S = 4$ SUE, ranging from $S=0$ (classical macrocell) to $S=10$ (only SCs—in this case, the MUE of interest becomes an SUE). Fig. 9 depicts the achievable rate (red line, left axis) and the total power consumption (blue line, right axis), whereas Fig. 10 shows the ASE. As is apparent, introducing SCs in the system has a *significant benefit in terms of all performance indicators*. Of course, this comparison does not account for the additional complexity and drawbacks introduced by increasing S (to mention a few, initial cost of network deployment and maintenance, and complexity of the system). However, although a suitable tradeoff needs to be sought, our

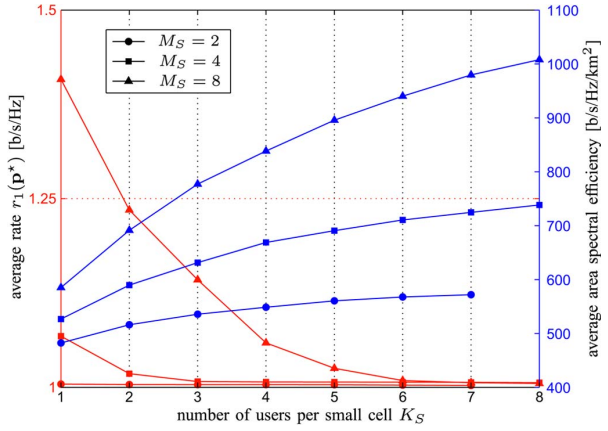


Fig. 11. Average rate (left axis) and average ASE (right axis) as functions of the number user per small cell. The average rate decreases with the number of users per small cell because of the MAI. However, the ASE is increasing with the the number of users per small cell. Moreover, increasing the number of receiving antennas at the SCA improves both, the average rate and average ASE.

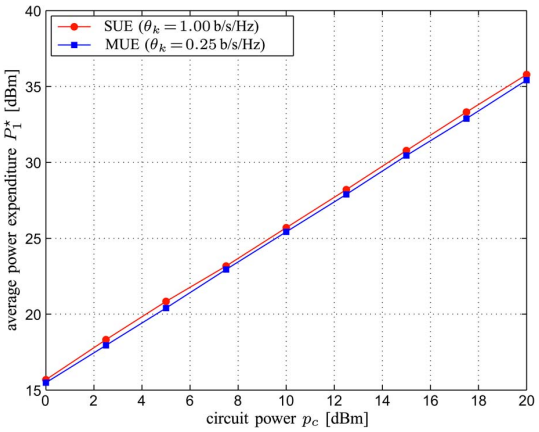


Fig. 12. Average power at the equilibrium as a function of the circuit power. The average power consumption scales linearly with the circuit power in the EE formulation.

analysis confirms that *network densification is one of the key technologies to meet 5G requirements* [28].

To verify the scalability of the proposed solution, we also investigate the impact of the number of receiving antennas at the SCA M_S. In Fig. 11, we plot the spectral efficiency (red lines, left axis) and the ASE (blue lines, right axis) as a function of the number of users per small cell K_S. Circular, squared, and triangular markers represent the cases for $M_S = \{2, 4, 8\}$ antennas at the SCA. The ASE is averaged over all users $K = K_0 + S \cdot K_S$, whereas the achievable rate is computed for an SUE of interest using $\theta_1 = 1$ b/s/Hz, averaging over 10^5 independent network realizations. As can be seen, increasing the number of antennas yields significant performance gains, thus representing a design parameter that can be exploited to boost the performance. Not only the spectral efficiency, as expected, benefits from increasing M_S (as an example, we can move from 500 b/s/Hz/km^2, achieved when using 2 antennas, to 1000 b/s/Hz/km^2, by increasing the number of receiving antennas up to 8, supporting $K = 60$ users), but *also does the EE*, confirming a recent result available in [33]: here, when $K_S = 7$, moving from $M_S = 2$ to 8 yields more than a 5-fold increase in the utility.

Finally, to evaluate the impact of the circuit power p_c on the EE of the system, we show in Fig. 12 the performance of the proposed algorithm as a function of p_c, averaged over 10^5 independent network realizations, where the red line refers to an SUE using $\theta_k = 1$ b/s/Hz, and the blue line refers to an MUE using $\theta_k = 0.25$ b/s/Hz. For all selected non-radiative powers $p_c \in [0, 20]$ dBm, the hypothesis $p_c \gg \sigma^2$ holds, which is in line with the state of the art for radio-frequency and baseband transceiver modeling [21]. As can be seen, the total power consumption at the equilibrium $P_1(\mathbf{p}^\star)$ is directly proportional to p_c. Put differently, the energy-efficient equilibrium point is highly impacted by the non-radiative power, and the bit-per-Joule metric suggests the use a radiative power which is comparable with the non-radiative one. Interestingly, the (normalized) achievable rates at equilibrium (not reported for concision) do not depend on p_c (1.1 and 0.6 b/s/Hz for SUE and MUE, respectively). This confirms a result which is well-known in the literature (e.g., see [23]): *EE increases as the circuit (non-radiative) power decreases*. Hence, reducing p_c, which is one of the main drivers in the device design further boosting the research in this field, can achieve a two-fold goal: not only is it expedient to reduce the constant power consumption (from an electronics point of view), but also it leads energy-aware terminals to reduce their radiative power when they aim at maximizing their bit-per-Joule performance (from an information-theoretic and resource-allocation perspective).

VI. Conclusion and Perspectives

In this paper, we proposed a distributed power allocation scheme for energy-aware, non-cooperative wireless users with minimum-rate constraints in the uplink of a multicarrier heterogenous network. The major challenge in this formulation is represented by the minimum-rate requirements that cast the problem into a non-cooperative game in the sense of Debreu, in which the actions sets of the players are coupled (and not independent as in the case of Nash-type games). We used fractional programming techniques to characterize the game's equilibrium states (when they exist) as the fixed points of a water-filling operator. To attain this equilibrium in a distributed fashion, we also proposed an adaptive, distributed algorithm based on an iterative water-filling best response process and we provided sufficient conditions for its convergence. The convergence and performance of the proposed solution were further assessed by numerical simulations: our results show that reducing the non-radiative power consumed by the user device electronics, offloading the macrocell traffic through small cells, and increasing the number of receive antennas, are critical to improve the performance of mobile terminals in terms of *both energy efficiency and spectral efficiency*. Using a realistic simulation setup, we showed that the proposed framework is able to achieve significantly high area spectral efficiencies (higher than

1000 b/s/Hz/km^2), peak and cell-edge spectral efficiencies (up to 6 b/s/Hz and around 0.5 b/s/Hz, respectively), and energy efficiencies (several Mb/J), while considering dense populations of users (around 1000 users/km^2), low power consumptions (at most a few Watts), a limited number of antennas (at most 8 for the small-cell access points and 16 for the macrocell base station), and simplified signal processing at the receiver (maximal ratio combining).

The system model adopted in this work encompasses a more general *multi-cellular* and *multi-tier* network, and the derived approach can be automatically adapted to such scenarios. Moreover, distinguishing features of the proposed distributed algorithm are its *scalability* and *flexibility*, which make it suitable for emerging 5G technologies [28], such as ultra-dense networks and massive MIMO.

Challenging open issues for further work include: *i*) assessing the feasibility of the problem given a particular network realization for the multicarrier case; *ii*) evaluating the impact of different receiver architectures (such as multiuser zero-forcing, and interference cancellation techniques) on the spectral and energy efficiency of the network; *iii*) accounting for highly time-varying scenarios in which users move around the network with different speeds.

Appendix A
Proof of Proposition 1

First, note that (8) can be expressed in the language of fractional programming as:

$$\mathbf{p}_k^\star = \underset{\mathbf{p}_k \in \mathcal{P}_k(\mathbf{p}_{-k})}{\arg\max} \ \frac{\varphi(\mathbf{p}_k)}{\chi(\mathbf{p}_k)} \tag{30}$$

where $\mathcal{P}_k(\mathbf{p}_{-k})$ is defined as in (9), and

$$\varphi(\mathbf{p}_k) = \sum_{n=1}^{N} \ln(1+\mu_{k,n} p_{k,n}) \text{ and } \chi(\mathbf{p}_k) = p_{c,k} + \sum_{n=1}^{N} p_{k,n}. \tag{31}$$

From [23, Sect. II.A] solving (30) is equivalent to finding the root of the following nonlinear function:

$$\Phi(\lambda_k) = \max_{\mathbf{p}_k \in \mathcal{P}_k(\mathbf{p}_{-k})} \varphi(\mathbf{p}_k) - \lambda_k \chi(\mathbf{p}_k) \tag{32}$$

where $\lambda_k \in \mathbb{R}$. To compute the solution of (30), let us first use (31), but without the constraint (8b), so that $\mathbf{p}_k \in \mathbb{R}_+^N$ (i.e., only nonnegative powers are considered). The stationarity condition, given by $\frac{\partial \varphi(\mathbf{p}_k)}{\partial p_{k,n}}|_{p_{k,n}=p_{k,n}^\star} - \lambda_k \frac{\partial \chi(\mathbf{p}_k)}{\partial p_{k,n}}|_{p_{k,n}=p_{k,n}^\star} = 0 \ \forall n$, using (31) becomes

$$\frac{\mu_{k,n}}{1+\mu_{k,n} p_{k,n}^\star} - \lambda_k = 0 \qquad \forall n. \tag{33}$$

Hence, considering $p_{k,n}^\star \geq 0$, the optimal power allocation becomes the WF criterion (17), in which the water level $\lambda_k^\star$ is replaced by λ_k. By plugging (33) back into (32) we can finally compute the optimal power level λ_k:

$$-\ln \lambda_k + (\beta_k - 1) = \alpha_k \lambda_k. \tag{34}$$

where the functions α_k and β_k are defined as in (20) and (21), respectively. To provide a better insight on (34), let us define $\nu_k = -\ln \lambda_k + (\beta_k - 1)$, so that (34) can be rewritten as $\nu_k e^{\nu_k} = \alpha_k e^{\beta_k - 1}$. Using the Lambert function $W(\cdot)$ we can obtain the expression of λ_k as in (19).

Introducing back the constraint (8b) simply places a lower bound on $\varphi(\mathbf{p}_k)$: $\varphi(\mathbf{p}_k) \geq \theta_k$. Following [23], this is equivalent to setting an upper bound $\overline{\lambda}_k$ on λ_k, that comes out of the IWF criterion that minimizes $\chi(\mathbf{p}_k)$ given $\varphi(\mathbf{p}_k) = \theta_k$, and is equal to (22). Hence, the solution to (8) is given by (17) with $\lambda_k^\star$ computed as in (18).

Appendix B
Proof of Proposition 2

There exists a unique DE $\mathbf{p}^\star$ if the best response map $\mathcal{B}(\mathbf{p}) = [\mathcal{B}_1(\mathbf{p}_{-1}), \ldots, \mathcal{B}_K(\mathbf{p}_{-K})]$ with $\mathcal{B}_k(\mathbf{p}_{-k}) = \arg\max_{\mathbf{p}_k \in \mathcal{P}_k(\mathbf{p}_{-k})} = u_k(\mathbf{p})$ is a contraction, i.e., there exists some $\varepsilon \in [0, 1)$ such that

$$\|\mathcal{B}(\mathbf{p}_1) - \mathcal{B}(\mathbf{p}_2)\| \leq \varepsilon \|\mathbf{p}_1 - \mathbf{p}_2\| \qquad \forall \mathbf{p}_1, \mathbf{p}_2 \in \mathcal{P}, \tag{35}$$

where $\mathcal{P} = \prod_{k=1}^K \mathcal{P}_k$. The nth component of user k's best response is given by $\mathcal{B}_{k,n}(\mathbf{p}^\star_{-k}) = [\mathcal{B}_k(\mathbf{p}^\star_{-k})]_n = p_{k,n}^\star$ as in (17). We begin by rewriting $\mu_{k,n}(\mathbf{p}_{-k,n})$ in (4) as follows:

$$\mu_{k,n}(\mathbf{p}_{-k,n}) = \frac{\omega_{kk,n}}{\sigma^2 + I_{k,n}} \tag{36}$$

where $I_{k,n} = \sum_{j \neq k} \omega_{kj,n} p_{j,n}$, and the quantities $\omega_{kj,n}$ are defined in (24) Using [29, Theorem 4] the DE $\mathbf{p}^\star$ is unique if, for any UE k,

$$\left\| \frac{\partial \mathbf{I}_k}{\partial \mathbf{p}_{-k}} \right\| \cdot \sup_{\mathbf{I}_k \in \mathbb{R}^N} \left\| \frac{\partial \mathcal{B}_k(\mathbf{p}_{-k})}{\partial \mathbf{I}_k} \right\| < 1 \tag{37}$$

with $\mathbf{I}_k = [I_{k,1}, \ldots, I_{k,N}]^T$. The first term of (37) is explicitly computed in [29, Eq. (19)] and it is equal to $\left\| \frac{\partial \mathbf{I}_k}{\partial \mathbf{p}_{-k}} \right\| = \sqrt{\sum_{j=1, j\neq k}^K \sum_{n=1}^N \omega_{kj,n}^2}$. As for the second term, we have:

$$\|\partial \mathcal{B}_k(\mathbf{p}_{-k}) / \partial \mathbf{I}_k\| = \sqrt{\sum_{\ell=1}^N \sum_{n=1}^N \left| \partial p_{k,n}^\star / \partial I_{k,\ell} \right|^2}, \tag{38}$$

where the optimal (best-responding) transmit power levels $p_{k,n}^\star$ are:

$$p_{k,n}^\star = (1/\lambda_k^\star - 1/\mu_{k,n}) \mathbb{1}_{\{\mu_{k,n} > \lambda_k^\star\}} \tag{39}$$

After some derivation steps, we obtain the norm of its partial derivative w.r.t. $I_{k,\ell}$ as follows:

$$\left| \frac{\partial p_{k,n}^\star}{\partial I_{k,\ell}} \right|^2 = \frac{\mathbb{1}_{\{\mu_{k,n} > \lambda_k^\star\}}}{\omega_{kk,\ell}^2 \left(\varsigma_k^\star\right)^2} \left[\xi_{k,\ell}^2 + \left(\left(\varsigma_k^\star\right)^2 - 2\varsigma_k^\star \xi_{k,\ell} \right) \mathbb{1}_{\{n=\ell\}} \right] \tag{40}$$

where, for convenience, we denote by $\varsigma_k^\star = |\mathcal{S}_k^\star|$ and

$$\xi_{k,\ell} = -\varsigma_k^\star \mu_{k,\ell}^2 \frac{\partial \left(1/\lambda_k^\star\right)}{\partial \mu_{k,\ell}}. \tag{41}$$

Summing over $n = 1, \ldots, N$ then yields:

$$\left\| \frac{\partial \mathcal{B}_k(\mathbf{p}_{-k})}{\partial \mathbf{I}_k} \right\| = \sqrt{\frac{1}{\varsigma_k^\star} \sum_{\ell \in \mathcal{S}_k^\star} \frac{1}{\omega_{kk,\ell}^2} \cdot \left(\xi_{k,\ell}^2 + \varsigma_k^\star - 2\xi_{k,\ell} \right)} \quad (42)$$

so that it remains to show that the terms $\xi_{k,\ell}$ in (41) are equivalent to (26) in Proposition 2 (see Appendix C). As a final step in the proof, notice that the function to be optimized in (23) depends only on $\mu_{k,n}$ which is an invertible, bijective function of $I_{k,n} \geq 0$ (since it is a strictly decreasing function with respect to $I_{k,n}$). Therefore, we can take the supremum over $\mu_{k,n} \in (0, \omega_{kk,n}^2/\sigma^2]$, $\forall n$ directly.

APPENDIX C

In this section, we compute $\xi_{k,\ell}$ in two different cases depending on the relative order between λ_k and $\overline{\lambda}_k$. Let us start from the minimum-rate WF criterion, in which UE k's water level is computed using (18). In this case, if $\mu_{k,\ell} > \overline{\lambda}_k$ (i.e., if $\ell \in \overline{\mathcal{S}}_k$),[5] we have $\overline{\lambda}_k^{-1} = \left(2^{N\theta_k} \prod_{n \in \overline{\mathcal{S}}_k} \mu_{k,n}^{-1}\right)^{1/\overline{\varsigma}_k} = \left(2^{N\theta_k} \prod_{n \in \overline{\mathcal{S}}_k, n \neq \ell} \mu_{k,n}^{-1}\right)^{1/\overline{\varsigma}_k} \mu_{k,\ell}^{-1/\overline{\varsigma}_k}$, where $\overline{\varsigma}_k = |\overline{\mathcal{S}}_k|$. From this, we get $\frac{\partial (1/\overline{\lambda}_k)}{\partial \mu_{k,\ell}} = -\frac{1}{\overline{\varsigma}_k \mu_{k,\ell} \overline{\lambda}_k}$, and thus, using (41), we finally obtain $\xi_{k,\ell} = \mu_{k,\ell}/\overline{\lambda}_k$, corresponding to the first subcase of (26).

Let us now focus on the energy-efficient WF, in which each UE k's water level is computed using (19). If $\mu_{k,\ell} > \lambda_k$, then:

$$\begin{aligned} \frac{\partial (1/\lambda_k)}{\partial \mu_{k,\ell}} &= \frac{1}{\lambda_k} \frac{\partial}{\partial \mu_{k,\ell}} \left[W\left(\alpha_k e^{\beta_k - 1}\right) - (\beta_k - 1) \right] \\ &= \frac{1}{\lambda_k} \left[\frac{\partial W\left(\alpha_k e^{\beta_k - 1}\right)}{\partial \mu_{k,\ell}} - \frac{\partial \beta_k}{\partial \mu_{k,\ell}} \right]. \end{aligned} \quad (43)$$

On one hand, using (20) and (21), we can compute the partial derivatives $\frac{\partial \alpha_k}{\partial \mu_{k,\ell}} = \frac{1}{\varsigma_k \mu_{k,\ell}^2}$ and $\frac{\partial \beta_k}{\partial \mu_{k,\ell}} = \frac{1}{\varsigma_k \mu_{k,\ell}}$, with $\varsigma_k = |\mathcal{S}_k|$. On the other hand, using the properties of the Lambert function, we get

$$\frac{\partial W\left(\alpha_k e^{\beta_k - 1}\right)}{\partial \mu_{k,\ell}} = \frac{W\left(\alpha_k e^{\beta_k - 1}\right) \cdot \frac{\partial}{\partial \mu_{k,\ell}} \left(\alpha_k e^{\beta_k - 1}\right)}{\left(\alpha_k e^{\beta_k - 1}\right) \left[1 + W\left(\alpha_k e^{\beta_k - 1}\right)\right]} \quad (44)$$

and hence:

$$\frac{\partial (1/\lambda_k)}{\partial \mu_{k,\ell}} = \frac{W\left(\alpha_k e^{\beta_k - 1}\right) - \alpha_k \mu_{k,\ell}}{\varsigma_k \mu_{k,\ell}^2 \lambda_k \alpha_k \left[1 + W\left(\alpha_k e^{\beta_k - 1}\right)\right]}. \quad (45)$$

Noting that, by inverting (19), $W\left(\alpha_k e^{\beta_k - 1}\right) = \beta_k - 1 - \ln \lambda_k$, and using simple mathematical steps, $\nu_k = -\ln \lambda_k + (\beta_k - 1)$ can be rewritten as $\nu_k = W\left(\alpha_k e^{\beta_k - 1}\right) = \alpha_k \lambda_k$. Using (41), $\xi_{k,\ell}$ corresponds to the second subcase of (26).

[5] Note that we are interested in computing $\xi_{k,\ell}$ only for $\ell \in \overline{\mathcal{S}}_k$, as in all other cases $\xi_{k,\ell} = 0$.

REFERENCES

[1] "The 1000x data challenge," Qualcomm, San Diego, CA, USA, Tech. Rep., 2013. [Online]. Available: http://www.qualcomm.com/1000x

[2] Green Touch Consortium, Tokyo, Japan, Tech. Rep. [Online]. Available: http://www.greentouch.org

[3] J. Hoydis, M. Kobayashi, and M. Debbah, "Green small-cell networks," *IEEE Veh. Technol. Mag.*, vol. 6, no. 1, pp. 37–43, Mar. 2011.

[4] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA, USA: MIT Press, 1991.

[5] W. Yu, G. Ginis, and J. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 5, pp. 1105–1115, Jun. 2002.

[6] R. Cendrillon, J. Huang, M. Chiang, and M. Moonen, "Autonomous spectrum balancing for digital subscriber lines," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4241–4257, Aug. 2007.

[7] F. Meshkati, H. V. Poor, and S. C. Schwartz, "An energy-efficient approach to power control and receiver design in wireless data networks," *IEEE Trans. Commun.*, vol. 53, no. 11, pp. 1885–1894, Nov. 2005.

[8] F. Meshkati, M. Chiang, H. V. Poor, and S. C. Schwartz, "A game-theoretic approach to energy-efficient power control in multicarrier CDMA systems," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 6, pp. 1115–1129, Jun. 2006.

[9] G. Scutari, D. Palomar, and S. Barbarossa, "Optimal linear precoding strategies for wideband noncooperative systems based on game theory part I: Nash equilibria," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1230–1249, Mar. 2008.

[10] G. Scutari, D. Palomar, and S. Barbarossa, "Competitive design of multiuser MIMO systems based on game theory: A unified view," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 7, pp. 1089–1103, Jul. 2008.

[11] P. Mertikopoulos, E. V. Belmega, and A. L. Moustakas, "Matrix exponential learning: Distributed optimization in MIMO systems," in *Proc. IEEE Int. Symp. Inf. Theory*, Cambridge, MA, USA, Jul. 2012, pp. 3028–3032.

[12] P. Coucheney, B. Gaujal, and P. Mertikopoulos, "Distributed optimization in multi-user MIMO systems with imperfect and delayed information," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jun./Jul. 2014, pp. 3097–3101.

[13] Y. Shi, J. Wang, K. Letaief, and R. Mallik, "A game-theoretic approach for distributed power control in interference relay channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 3151–3161, Jun. 2009.

[14] S. Ren and M. van der Schaar, "Distributed power allocation in multiuser multi-channel cellular relay networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 6, pp. 1952–1964, Sep. 2010.

[15] D. T. Ngo, L. B. Le, T. Le-Ngoc, E. Hossain, and D. I. Kim, "Distributed interference management in two-tier CDMA femtocell networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 979–989, Mar. 2012.

[16] G. Scutari, D. Palomar, F. Facchinei, and J.-S. Pang, "Convex optimization, game theory, and variational inequality theory," *IEEE Signal Process. Mag.*, vol. 27, no. 3,pp. 35–49, Mar. 2010.

[17] J.-S. Pang, G. Scutari, D. Palomar, and F. Facchinei, "Design of cognitive radio systems under temperature-interference constraints: A variational inequality approach," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3251–3271, Jun. 2010.

[18] I. Stupia, L. Sanguinetti, G. Bacci, and L. Vandendorpe, "Power control in networks with heterogeneous users: A quasi-variational inequality approach," *IEEE Trans. Signal Process.*, submitted for publication, 2014. [Online]. Available: http://arxiv.orglabs/1308.4840

[19] E. V. Belmega and S. Lasaulce, "Energy-efficient precoding for multiple antenna terminals," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 329–340, Jan. 2011.

[20] E. V. Belmega, S. Lasaulce, and M. Debbah, "A survey on energyefficient communications," in *Proc. IEEE Int. Symp. PIMRC*, Istanbul, Turkey, Sep. 2010, pp. 289–294.

[21] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Designing multi user MIMO for energy efficiency: When is massive MIMO the answer?" in *Proc. IEEE WCNC*, Istanbul, Turkey, Apr. 2014, pp. 242–247.

[22] G. Bacci, E. V. Belmega, and L. Sanguinetti, "Distributed energy efficient power optimization in cellular relay networks with minimum rate constraints," in *Proc. IEEE ICASSP*, Florence, Italy, May 2014, pp. 7014–7018.

[23] C. Isheden, Z. Chong, E. Jorswieck, and G. Fettweis, "Framework for link level energy efficiency optimization with informed transmitter," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2946–2957, Aug. 2012.

[24] J. Nash, "Non-cooperative games," *Ann. Math.*, vol. 54, no. 2, pp. 286–295, 1951.

[25] G. Debreu, "A social equilibrium existence theorem," *Proc. Natl. Acad. Sci. USA*, vol. 38, no. 10, pp. 886–893, Oct. 1952.

[26] F. Facchinei and C. Kanzow, "Generalized nash equilibrium problems," *Quart. J. Oper. Res.*, vol. 5, no. 3, pp. 173–210, Sep. 2007.

[27] S. Boyd and L. Vandenberghe, Eds., *Convex Optimization*. Cambridge, UK: Cambridge Univ. Press, 2004.
[28] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.
[29] G. Miao, N. Himayat, G. Li, and S. Talwar, "Distributed interference aware energy-efficient power optimization," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1323–1333, Apr. 2011.
[30] J.-S. Pang, G. Scutari, F. Facchinei, and C. Wang, "Distributed power allocation with rate constraints in Gaussian parallel interference channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3471–3489, Aug. 2008.
[31] G. Bacci, E. V. Belmega, and L. Sanguinetti, "Distributed energyefficient power and subcarrier allocation for OFDMA-based small cells," in *Proc. IEEE ICC Workshop Small Cell 5G Netw.*, Sydney, NSW, Australia, Jun. 2014, pp. 647–652.
[32] R. Corless, G. Gonnet, D. Hare, D. Jeffrey, and D. Knuth, "On the Lambert W function," *Adv. Comput. Math.*, vol. 5, pp. 329–359, 1996.
[33] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059–3075, Jun. 2015, DOI: 10.1109/TWC.2015.2400437 .
[34] G. Miao, N. Himayat, and G. Li, "Energy-efficient link adaptation in frequency-selective channels," *IEEE Trans. Commun.*, vol. 58, no. 2, pp. 545–554, Feb. 2010.
[35] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.
[36] H. Tabassum, Z. Dawy, and M. S. Alouini, "Sum rate maximization in the uplink of multi-cell OFDMA networks," in *Proc. Int. Wireless Commun. Mobile Comput. Conf.*, Istanbul, Turkey, Jul. 2011, pp. 1152–1157.
[37] M. Chiang, P. Hande, T. Lan, and C. W. Tan, "Power control in wireless cellular networks," *Found. Trends Netw.*, vol. 2, no. 4, pp. 381–533, 2007.
[38] G. Bacci, E. V. Belmega, P. Mertikopoulos, and L. Sanguinetti, "Energy-aware competitive link adaptation in heterogeneous networks," Univ. Pisa, Pisa, Italy, Tech. Rep., Jul. 2014. [Online]. Available: http://www.iet.unipi.it/l.sanguinetti/hetnets.pdf
[39] G. Bacci, L. Sanguinetti, M. Luise, and H. V. Poor, "Energy-efficient power control for contention-based synchronization in OFDMA systems with discrete powers and limited feedback," *EURASIP J. Wireless Commun. Netw.*, vol. 2013, no. 1, p. 192, Jul. 2013.
[40] W. Dinkelbach, "On nonlinear fractional programming," *Manage. Sci.*, vol. 13, no. 7, pp. 492–498, Mar. 1967.
[41] I. Stupia, L. Vandendorpe, L. Sanguinetti, and G. Bacci, "Distributed energy-efficient power optimization for relay-aided heterogeneous networks," in *Proc. Int. Workshop Wireless Netw., Commun., Coop. Competition*, Hammamet, Tunisia, pp. 563–568, May 2014.
[42] G. Calcev *et al.*, "A wideband spatial channel model for system-wide simulations," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 389–403, Mar. 2007.
[43] "LTE; Requirements for further advancements for Evolved Universal Terrestrial Radio Access (EUTRA) (LTE-Advanced)," 3GPP Technical Specification Group, Sophia Antipolis Cedex, France, Tech. Rep. 3GPP TR 36.913 v10.0.0, Apr. 2011.

Giacomo Bacci (S'07–M'09) received the B.E. and the M.E. degrees in telecommunications engineering and the Ph.D. degree in information engineering from the University of Pisa, Pisa, Italy, in 2002, 2004, and 2008, respectively. In 2006–2007, he was a Visiting Student Research Collaborator at Princeton University, Princeton, NJ, USA. In 2008-2014, he was a Post-doctoral Research Fellow at the University of Pisa. In 2008-2012, he also joined Wiser srl, Livorno, Italy, as a Software Engineer, and in 2012–14 he was also enrolled as a Visiting Post-doctoral Research Associate at Princeton University. Since 2015, he is a Product Manager for interactive satellite broadband communications at MBI srl, Pisa, Italy. He is the recipient of the FP7 Marie Curie International Outgoing Fellowships for career development (IOF) 2011 GRAND-CRU, and received the Best Paper Award from the IEEE Wireless Communications and Networking Conference (WCNC) 2013, the Best Student Paper Award from the International Waveform Diversity and Design Conference (WDD) 2007, and the Best Session Paper at the ESA Workshop on EGNOS Performance and Applications 2005. He is also the recipient of the 2014 URSI Young Scientist Award, and he was named an Exemplary Reviewer 2012 for IEEE WIRELESS COMMUNICATIONS LETTERS.

E. Veronica Belmega (S'07–M'10) received the B.Sc. degree from the University Politehnica of Bucharest, Romania, in 2007, and the M.Sc. and Ph.D. degrees from the Université Paris-Sud 11, Paris, France, in 2007 and 2010, respectively. In 2010–2011, she was a Post-doctoral Researcher in a joint project between the Alcatel-Lucent Chair on Flexible Radio in Supélec and Princeton University. She is currently an Assistant Professor at ETIS/ENSEA Université de Cergy-Pontoise CNRS, Cergy-Pontoise, France. She was one of the ten recipients of the 2009 L'Oréal French Academy of Science Fellowship: "For young women doctoral candidates in science." Her interests lie in resource allocation in wireless communications, state estimation in electrical power grid using tools from game theory and distributed optimization.

Panayotis Mertikopoulos (M'11) received the B.S. degree in physics with major in astrophysics from the University of Athens in 2003. He received the M.Sc. and M.Phil. degrees in mathematics from Brown University in 2005 and 2006, respectively, and the Ph.D. degree in physics from the University of Athens in 2010. In 2010–2011, he was a post-doctoral researcher at the Economics Department of École Polytechnique, Paris, France. Since 2011, he has been a CNRS Researcher at the Laboratoire d'Informatique de Grenoble, Grenoble, France, and has a joint research affiliation with Inria, France. He has been an Embeirikeion Foundation Fellow since 2003 and was the recipient of the best paper award at NetGCoop'12. He was a track co-chair of IFORS 2014, TPC co-chair of WiOpt 2014, and general co-chair of AlgoGT 2013. His main research interests lie in dynamical systems, optimization, game theory, and their applications to wireless networks.

Luca Sanguinetti (S'04–M'06) received the Laurea Telecommunications Engineer degree (cum laude) and the Ph.D. degree in information engineering from the University of Pisa, Italy, in 2002 and 2005, respectively. Since 2005, he has been with the Dipartimento di Ingegneria dell'Informazione, University of Pisa. In 2004, he was a visiting Ph.D. student at the German Aerospace Center (DLR), Oberpfaffenhofen, Germany. During the period June 2007-June 2008, he was a Postdoctoral Associate in the Deptartment of Electrical Engineering, Princeton University. During the period June 2010–September 2010, he was selected for a research assistantship at the Technische Universitat Munchen. From July 2013 to July 2015, he was with the Alcatel-Lucent Chair on Flexible Radio, Supélec, Gif-sur-Yvette, France. He is an Assistant Professor at the Dipartimento di Ingegneria dell'Informazione, University of Pisa. He is currently serving as an Associate Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and IEEE JOURNAL OF SELECTED AREAS OF COMMUNICATIONS (series on Green Communications and Networking). His expertise and general interests span the areas of communications and signal processing, game theory and random matrix theory for wireless communications. He was the co-recipient of two best paper awards: IEEE Wireless Communications and Networking Conference (WCNC) 2013 and IEEE Wireless Communications and Networking Conference (WCNC) 2014. He is also the recipient of the FP7 Marie Curie IEF 2013 "Dense deployments for green cellular networks."

Protecting Secret Key Generation Systems Against Jamming: Energy Harvesting and Channel Hopping Approaches

E. Veronica Belmega, *Member, IEEE*, and Arsenia Chorti, *Member, IEEE*

***Abstract*— Jamming attacks represent a critical vulnerability for wireless secret key generation (SKG) systems. In this paper, two counter-jamming approaches are investigated for SKG systems: first, the employment of energy harvesting (EH) at the legitimate nodes to turn part of the jamming power into useful communication power, and, second, the use of channel hopping or power spreading in block fading channels to reduce the impact of jamming. In both cases, the adversarial interaction between the pair of legitimate nodes and the jammer is formulated as a two-player zero-sum game and the Nash and Stackelberg equilibria are characterized analytically and in closed form. In particular, in the case of EH receivers, the existence of a critical transmission power for the legitimate nodes allows the full characterization of the game's equilibria and also enables the complete neutralization of the jammer. In the case of channel hopping versus power spreading techniques, it is shown that the jammer's optimal strategy is always power spreading while the legitimate nodes should only use power spreading in the high signal-to-interference ratio (SIR) regime. In the low SIR regime, when avoiding the jammer's interference becomes critical, channel hopping is optimal for the legitimate nodes. Numerical results demonstrate the efficiency of both counter-jamming measures.**

***Index Terms*— Secret key generation, jamming, energy harvesting, channel hopping, zero-sum game.**

I. Introduction

SECRET key generation (SKG) from shared randomness at two remote locations has been extensively studied [3]–[10] and has recently been extended to unauthenticated channels [11], [12]. SKG techniques have also been incorporated in protocols that are resilient to spoofing, tampering and man-in-the-middle active attacks [13]. Still, such key generation techniques are not entirely robust against active adversaries, particularly during the advantage distillation phase. Denial of service attacks in the form of jamming are a known vulnerability of SKG systems; in [14], it was demonstrated that when increasing the jamming power, the reconciliation rate normalized to the rate of the SKG increases sharply and the SKG process can in essence be brought to a halt. As SKG techniques are currently being considered for applications such as the Internet of things (IoT) [15], the study of appropriate counter-jamming approaches is timely.

Typically, jamming in wireless communication systems has been investigated using game theoretic tools [16]–[23]. Contrary to our work, these earlier studies focus on performance metrics that are either based on the legitimate nodes' signal-to-interference-plus-noise ratio (SINR) [16]–[21] and do not incorporate physical-layer security constraints at all, or are based on the secrecy capacity [22], [23]. The secrecy capacity is inherently different than the SKG capacity considered in this work; the former measures the maximum rate at which both confidential and reliable communication is possible, while the latter represents the maximum rate at which a common secret key can be extracted from the observation of correlated sequences at two remote locations [24].

In the past, two main counter-jamming approaches have been commonly considered: direct sequence spread spectrum (DSSS) and frequency hopping spread spectrum (FHSS) [25], [26]. In either approach, the impact of power constrained jammers can be limited because their optimal strategy has been proved to be the spreading of their available power over the entire bandwidth (and thus jam with potentially low power). However, DSSS and FHSS systems require a pre-shared secret to establish the spreading sequence or the hopping pattern at Alice and Bob; as such, they are not directly applicable to SKG systems that on the contrary *seek to establish* a secret key. Attempting to resolve this contradiction and reconcile DSSS and FHSS with SKG, uncoordinated frequency hopping and spreading techniques have recently been investigated in [27] and [28]. The main idea behind the proposed approaches was the randomization of the selection of the hopping/spreading sequences, at the cost of reducing the achievable rates for secret key establishment.

However, in uncoordinated hopping/spreading techniques there are minimum requirements regarding the length of the pseudorandom sequences employed. As a result, accounting for the strict bandwidth specifications of fourth and fifth generation networks, the use of long pseudorandom sequences can be a limiting factor. Thus, investigating different counter-jamming approaches based on the use of channel hopping or power spreading over multiple orthogonal subcarriers, e.g., orthogonal frequency division multiplexing (OFDM)

Manuscript received November 16, 2016; revised March 17, 2017 and April 24, 2017; accepted May 19, 2017. Date of publication June 7, 2017; date of current version July 26, 2017. This work was supported in part by ENSEA, Cergy-Pontoise, France, and in part by LABEX MME-DII. This paper was presented at the Proceedings IEEE International Conference on Communications 2017 [1], [2]. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Qian Wang. *(Corresponding author: E. Veronica Belmega.)*

E. V. Belmega is with ETIS, UMR 8051, Université Paris Seine, Université Cergy-Pointoise, ENSEA, CNRS, France, and also with Inria (e-mail: belmega@ensea.fr).

A. Chorti is with the School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, U.K. (e-mail: achorti@essex.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2017.2713342

systems [16], [18], is timely and offers an interesting alternative to [27] and [28] as in OFDM systems there is no need for coordination of the remote nodes. Furthermore, although in [27] and [28] the numerical investigations focused on the throughput, a Media Access Control (MAC) layer quantity, when analyzing physical layer security SKG systems the standard approach is to utilize the SKG capacity (a physical layer quantity).

On a different note, next generation terminals are likely to be enhanced with many new features that could prove pivotal in protecting against jamming. For example, greater energy autonomy exploiting energy harvesting (EH) approaches [29], [30] is being researched for systems such as wireless sensor networks for IoT applications. Thus, it is interesting to investigate whether EH could be utilized as a counter-jamming technique by exploiting the harvested jamming power to enhance the quality of the legitimate communication.

Motivated by the above, in the present work we propose two novel approaches for alleviating the impact of jamming in SKG systems. In both approaches, we model the interaction between the legitimate nodes and the adversarial jammer as a two-player zero-sum game in which the SKG capacity plays the role of the utility function. We investigate two non-cooperative solutions: the Nash equilibria (NE), when both players make their decision simultaneously and the Stackelberg equilibria (SE), when the legitimate nodes have an advantage and choose their strategy first while anticipating the jammer's response.

In the first part of this contribution, we study systems in which the legitimate nodes are equipped with EH capabilities and examine whether this added functionality is useful in preempting jamming attacks. We focus on time switching EH protocols [30]: for a fraction of time the legitimate nodes operate in EH mode and switch to the SKG procedure for the rest. To the best of our knowledge, this is among the first works to investigate EH as a counter-jamming approach with the exception of [21].[1]

Our analysis reveals the existence of a critical power thres-hold p_{th} for the legitimate nodes and of an associated threshold harvesting duration τ_{th}. When the legitimate nodes employ EH for longer than τ_{th}, the attacker's optimal strategy is not to jam at all, i.e., the jammer is effectively neutralized. However, neutralizing the jammer is not a stable solution to unilateral deviations (if the strategic decisions are taken simultaneously) and is therefore not a Nash equilibrium (NE) of the game. At the NE, it is found that both the legitimate nodes and the jammer transmit with full power and that the EH duration does not correspond always to the above threshold. At low signal to interference ratio (SIR) (e.g., relatively low transmit power or high jamming power), the EH optimal duration equals τ_{th}. Although the attacker jams with full power, the power collected from EH cancels out the impact of the attack and the SKG capacity is equivalent to the case of using EH for the same duration in absence of a jammer. At medium to high SIR, the EH optimal duration becomes lower than τ_{th} and decreases until the legitimate nodes do not harvest energy at all.

Furthermore, when moving to a hierarchical game formulation, the SE analysis reveals that the legitimate nodes should play the NE strategy. Whenever the legitimate nodes' harvest energy for a duration τ_{th} (at the NE), the jammer neutralization strategy is also a SE solution. This means that, in a hierarchical game, the jammer can potentially be deterred from launching the attack.

In the second part of this investigation, extending the studies in [19] and [21] to SKG systems, counter-jamming policies are investigated for N block fading additive white Gaussian noise (BF AWGN) channels, e.g., systems with N orthogonal subcarriers. At the NE, the jammer always spreads its power over all subcarriers, while for the legitimate nodes the optimality of channel hopping or power spreading depends on the channel parameters. In the high SIR regime, the legitimate nodes should use power spreading to exploit the entire available spectrum given the relatively low jamming interference. On the other hand, at low SIR, the legitimate nodes should use channel hopping and transmit over a single subcarrier to avoid most of the jammer's interference. Furthermore, in characterizing the game's SE we find that the optimal SE strategies reduce to the NE ones, demonstrating that there is no extra payoff to be earned from the advantage of playing first.

Preliminary results of this work have been presented in [1] and [2]. The major contributions and improvements of this journal paper as compared with [1] and [2] consist in: providing complete proofs of all the results regarding the NE analysis and the jammer neutralization state; relaxing the action set of the jammer, in the energy harvesting case, from the discrete choice between remaining silent and transmitting at full power into the continuous interval of all possible power values, which has brought to light the existence of additional NEs; providing the additional analysis of the Stackelberg equilibrium; providing a comparative discussion between the two counter-jamming methods in Sec. V-C.

The paper is organized as follows. In Sec. II, the SKG baseline system model is introduced. In Sec. III, the adversarial interaction between the EH legitimate nodes and the jammer is formulated and analyzed using a zero-sum non-cooperative game framework, while in Sec. IV this setting is used to study channel hopping vs. power spreading in BF AWGN systems. Numerical illustrations and a detailed discussion of these counter-jamming strategies are provided in Sec. V, while the conclusions are given in Sec. VI.

[1]The recent work [21] proposes to harvest energy from the jamming interference in a multi-user interference channel in which the jammer is not a strategic decision maker. In terms of formulation, a global optimization problem is investigated (as opposed to an adversarial game). Furthermore, the global performance metric in [21] does not incorporate security constraints and the harvested energy is not directly exploited in the communication phase, appearing only as an additional term in the utility function.

II. SKG System Model in the Presence of a Jammer

The baseline SKG system model with two legitimate nodes, denoted by Alice and Bob and a single adversary, denoted by Eve, is depicted in Fig. 1. Typically, the SKG process consists of three phases [4], [6]. In the first phase, referred

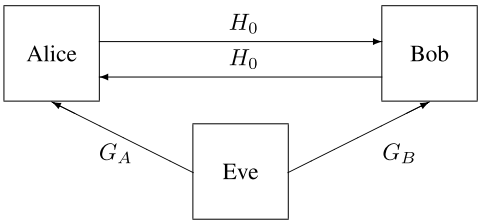


Fig. 1. SKG system model with two legitimate nodes and a single adversary.

to as *shared randomness distillation*, Alice and Bob observe dependent random variables denoted by Y_A, Y_B while an eavesdropper, referred to as Eve, observes Y_E. In wireless channels, a readily available source of shared randomness is the multipath fading due to the reciprocity of the wireless medium during the channel's coherence time [10]. Here, we focus exclusively on shared randomness extraction from Rayleigh fading coefficients.

In the next two phases, known as *information reconciliation* and *privacy amplification*, side information V is exchanged between Alice and Bob, generated by corresponding encoders f_A, f_B. At the end of the SKG process, a common key $K \in \mathcal{K}$ is extracted at Alice and Bob such that, for any $\epsilon > 0$, the following statements hold [8]:

$$Pr\left(K = f_A\left(Y_A, V\right) = f_B\left(Y_B, V\right)\right) \geq 1 - \epsilon, \quad (1)$$

$$I(K; V) \leq \epsilon, \quad (2)$$

$$H(K) \geq \log|\mathcal{K}| - \epsilon, \quad (3)$$

where $H(K)$ denotes the entropy of the key K and $I(K; V)$ denotes the mutual information between K and V.

The first inequality demonstrates that the SKG process can be made error free; (2) ensures that the exchange of side information through public discussion does not leak any information to eavesdroppers; while (3) establishes that the generated keys attain maximum entropy (i.e., are uniform). Under the three conditions, an upper bound on the rate for the generation of secret keys is given by $\min\{I(Y_A; Y_B), I(Y_A; Y_B|Y_E)\}$ [3], [4]. Assuming rich multipath environments, the decorrelation properties of the wireless channel over short distances can be exploited to ensure that Eve's observation Y_E is uncorrelated with Y_A and Y_B [7], [8], [10]; in this case, the SKG capacity is given by [3, Sec. II]

$$C = I(Y_A; Y_B). \quad (4)$$

We assume that this holds true in the rest of this study and consider the SKG capacity above to be the focal performance metric.

SKG in Rayleigh fading channels has been extensively analyzed, e.g., [7], [8]. In these works, it was assumed that Alice and Bob exchange unit probe signals to excite the fading channel and obtain respective observations Y_A and Y_B with

$$Y_A = H_0 + Z_A, \;\; Y_B = H_0 + Z_B,$$

where H_0 denotes the fading coefficient in the link between the legitimate nodes, modeled as a zero mean Gaussian random variable $H_0 \sim \mathcal{N}(0, \sigma_H^2)$, and, Z_A and Z_B model the effect of AWGN and denote independent and identically distributed (i.i.d.) Gaussian random variables $Z_A \sim \mathcal{N}(0, N_A)$, $Z_B \sim \mathcal{N}(0, N_B)$. Using this notation, the SKG capacity has been expressed as [8]:

$$C = I(Y_A; Y_B) = \frac{1}{2}\log_2\left(1 + \frac{\sigma_H^2}{N_A + N_B + \frac{N_A N_B}{\sigma_H^2}}\right). \quad (5)$$

In this work, we assume that Eve is no longer a passive eavesdropper but a malicious jammer. To include jamming attacks in the above model, we consider the following extension:

$$Y_A = \sqrt{p}H_0 + \sqrt{\gamma}G_A + Z_A, \quad (6)$$

$$Y_B = \sqrt{p}H_0 + \sqrt{\gamma}G_B + Z_B, \quad (7)$$

assuming that Alice and Bob exchange constant probe signals [8] with power $p \leq P$ and that Eve transmits constant jamming signals [14] with power $\gamma \leq \Gamma$. The fading coefficient in the link between Eve and Alice is denoted by $G_A \sim \mathcal{N}\left(0, \sigma_A^2\right)$ and in the link between Eve and Bob by $G_B \sim \mathcal{N}\left(0, \sigma_B^2\right)$. For simplicity and without loss of generality, the noise variables Z_A and Z_B are assumed to have unit variance, i.e., are modeled as i.i.d. Gaussian random variables $Z_A, \; Z_B \sim \mathcal{N}(0, 1)$.

Under these assumptions, a simple calculation reveals that the SKG capacity can be expressed as a function of p and γ:

$$C(p, \gamma) = \frac{1}{2}\log_2\left(1 + \frac{\sigma_H^2 p}{2 + (\sigma_A^2 + \sigma_B^2)\gamma + \frac{(1+\sigma_A^2\gamma)(1+\sigma_B^2\gamma)}{\sigma_H^2 p}}\right). \quad (8)$$

By inspecting the first-order derivatives of (8), we conclude that $C(p, \gamma)$ is a strictly increasing function of p for any fixed γ, and a strictly decreasing function of γ for any fixed p. This implies that the legitimate nodes will transmit at full power P to maximize the SKG capacity, whereas the jammer will also transmit with full power Γ to minimize the SKG capacity. Also, it is a strictly convex function with respect to (w.r.t.) γ for any fixed $p > 0$ as its second derivative w.r.t. γ is strictly positive.

III. Energy Harvesting Against Jamming

In order to study EH as a counter-jamming measure, we focus on a time-switching EH scheme [30], i.e., we assume that each transmission interval of duration T is divided in two parts. In the first period of duration τT ($0 < \tau \leq 1$ being the fraction of T dedicated to EH), both Alice and Bob operate in EH mode with efficiency $0 < \zeta \leq 1$; in the second period of duration $(1 - \tau)T$, the legitimate nodes operate in SKG mode using the overall available power (including harvested power). For simplicity, we assume that the energy harvested can be stored in a battery without any overflowing issues (unlimited storage) [31].

Furthermore, to ease the mathematical derivation and by ensuring symmetry in the energy harvested at Alice and Bob we assume that $\sigma_A^2 = \sigma_B^2 = \sigma^2$ (the Eve-Alice and Eve-Bob links have equal variance). Given the above considerations and

assuming that the energy harvested by Alice and Bob is linear in the received RF power [30]:

$$E = \zeta \tau T \gamma \sigma^2, \tag{9}$$

the harvested power for each legitimate node per transmission interval can be expressed as

$$p^{EH} = \frac{E}{(1-\tau)T} = \kappa \gamma, \tag{10}$$

where $\kappa = \frac{\zeta \tau \sigma^2}{1-\tau}$ is a convex increasing function of τ. Thus, the SKG capacity is given by:

$$\tilde{u}(p, \tau, \gamma) = \frac{1-\tau}{2} \log_2 \left(1 + \frac{\left(\frac{p}{1-\tau} + \kappa\gamma\right)\sigma_H^2}{2(1+\sigma^2\gamma) + \frac{(1+\sigma^2\gamma)^2}{\left(\frac{p}{1-\tau}+\kappa\gamma\right)\sigma_H^2}} \right), \tag{11}$$

with power constraints $p \leq P$, $\gamma \leq \Gamma$.

A simple inspection of (11) reveals that this scenario is a generalization of the standard SKG setting. Indeed, if the legitimate nodes decide not to harvest energy, i.e., $\tau = 0$, (8) is obtained for $\sigma_A^2 = \sigma_B^2 = \sigma^2, N_A = N_B = 1$. In the model with EH, the legitimate nodes can maximize $\tilde{u}$ by tuning the additional variable τ. However, it is no longer straightforward that the jammer should transmit with the maximum available power as $\tilde{u}(p, \gamma, \tau)$ is no longer monotonically decreasing in γ.

Non-cooperative game theory provides the natural framework to study the adversarial interaction between the legitimate nodes and the jammer. Although game theory has already been exploited in physical layer security problems, e.g. [22], [23], to the best of our knowledge, this work is among the first to investigate EH as an effective means to counteract on jamming attacks.

A. Jammer Neutralization

Before introducing the game framework, we make two important observations regarding the SKG utility in (11) and discuss their implications.

Remark 1: For any fixed τ and γ, $\tilde{u}(p, \tau, \gamma)$ is monotonically increasing in p and

$$\arg \max_{p \in [0,P]} \tilde{u}(p, \tau, \gamma) = P. \tag{12}$$

Remark 2: For any fixed p and τ, $\tilde{u}(p, \tau, \gamma)$ is monotone in γ. In particular, it is monotonically decreasing in γ if $p > p_{th}(\tau) \triangleq \zeta\tau$, a constant if $p = p_{th}(\tau)$, and monotonically increasing if $p < p_{th}(\tau)$. This implies that:

$$\arg \min_{\gamma \in [0,\Gamma]} \tilde{u}(p, \tau, \gamma) = 0, \quad \text{if } p < p_{th}(\tau) \tag{13}$$

$$\arg \min_{\gamma \in [0,\Gamma]} \tilde{u}(p, \tau, \gamma) \in [0, \Gamma], \quad \text{if } p = p_{th}(\tau) \tag{14}$$

$$\arg \min_{\gamma \in [0,\Gamma]} \tilde{u}(p, \tau, \gamma) = \Gamma, \quad \text{if } p > p_{th}(\tau). \tag{15}$$

Remark 1 shows that, to maximize the utility, the legitimate nodes should transmit at maximum power P. On the contrary, Remark 2 shows that the jammer should practically switch in between staying silent, i.e., $\gamma = 0$, and jamming at full power, i.e., $\gamma = \Gamma$, depending on the choice (p, τ) of the legitimate nodes.

Remark 2 reveals that the legitimate nodes can neutralize the jammer by transmitting at a relatively low power $p < p_{th}(\tau)$. Although this result may seem counter-intuitive at first, this condition is equivalent to $\tau > \tau_{th}(p) \triangleq \frac{p}{\zeta}$, which means that the legitimate nodes spend a relatively large proportion of time harvesting the jamming interference before actually transmitting. In other words, the jammer is forced to stay silent since the harm it can cause by interfering in the SKG phase is overcome by the harvested energy in the EH phase. This novel result shows that the jamming interference, which is commonly thought as being harmful to the legitimate communication, can be exploited and transformed into useful power via EH. If Alice and Bob transmit with exactly $p_{th}(\tau)$, the jammer becomes indifferent between all its choices $\gamma \in [0, \Gamma]$ and has no interest in actively jamming the transmission.

The necessary conditions for the jammer neutralization are formalized below.

Proposition 1: The optimal strategy for the legitimate nodes that maximizes the SKG utility while ensuring that the jammer has no interest in actively jamming the transmission is given by:

$$p^{NJ} = \min \{P, p_{th}(\tau^*)\} \quad \textit{and} \quad \tau^{NJ} = \min \{\tau_{th}(P), \tau^*\}, \tag{16}$$

where $\tau^* \in (0, 1)$ is the unique maximizer of $\tilde{u}(p_{th}(\tau), \tau, 0)$ w.r.t. τ.

For the detailed proof the reader is referred to Appendix VI-A. Notice that, if the jammer stays silent $\gamma = 0$, there is no actual energy harvested during the EH phase of duration τ^{NJ}. Rather, the legitimate nodes' choice to use EH for a fraction of time τ^{NJ} acts as an effective threat to ensure the jammer has no interest in actively jamming the transmission. However, neutralizing the jammer may not be the overall optimal strategy for the legitimate nodes. A hint for this is that whenever $\tau^{NJ} = \tau^* < \tau_{th}(P)$, the transmit power is $p^{NJ} < P$, which we know is not optimal from Remark 1.

B. Game Formulation and Nash Equilibria

The interaction between the legitimate nodes and the jammer is formalized as a two-player zero-sum game, defined as the tuple $\tilde{\mathcal{G}} = \{\tilde{\mathcal{A}}_L, \tilde{\mathcal{A}}_J, \tilde{u}(p, \tau, \gamma)\}$ in which the players are: player L representing the legitimate nodes (Alice and Bob act as a single player) on one side, and player J, the jammer, on the other. The action (p, τ) of player L lies in the set $\tilde{\mathcal{A}}_L = [0, P] \times [0, 1]$, and the action γ of player J lies in the set $\tilde{\mathcal{A}}_J = [0, \Gamma]$. The objective of player L is to maximize the SKG utility $\tilde{u}(p, \tau, \gamma)$ given in (11), whereas player J aims at minimizing it.

The two players are adversaries and the optimal strategy of one player depends on the choice of their opponent and cannot be determined unilaterally. In such interactive situations, the NE [32] is the natural solution concept. Intuitively, a profile $(p^{NE}, \tau^{NE}, \gamma^{NE}) \in \tilde{\mathcal{A}}_L \times \tilde{\mathcal{A}}_J$ is a NE if none of the players can benefit by deviating from this profile knowing

that their opponent plays accordingly. Hence, NEs are system states that are stable to unilateral deviations.

We can easily check that the state $(p^{NJ}, \tau^{NJ}, 0)$ is not a NE since the legitimate nodes gain by deviating from it. Knowing that the jammer stays silent, player L can increase the SKG utility by deviating to $\tau = 0$. Using the whole duration T in SKG mode increases the utility when no energy is harvested in the EH phase. This, in turn, will cause also the jammer to deviate from $\gamma = 0$ and actively jam the transmission.

Theorem 1 shows that the game $\tilde{\mathcal{G}}$ has at least one NE at which both players transmit with maximum power. This NE may be unique or not, depending on the system parameters.

Theorem 1: The game $\tilde{\mathcal{G}}$ has at least one NE. Moreover, the profile (P, τ^{NE}, Γ) is a NE solution such that the EH strategy is either $\tau^{NE} = 0$ or $\tau^{NE} = \min\{\tau_{th}(P), \tau_{max}\}$ with $\tau_{th}(P) = \frac{P}{\zeta}$ and $\tau_{max} \in (0, 1)$ representing the critical maximum point of $\tilde{u}(P, \tau, \Gamma)$ w.r.t. τ, depending on the system parameters. If $\tau^{NE} < \tau_{th}(P)$, then the profile (P, τ^{NE}, Γ) is the unique NE of the game almost surely.

The proof is detailed in Appendix VI-B. We observe that, at the NE above (P, τ^{NE}, Γ) and depending on the system parameters, player L may harvest energy for a fraction of time $\tau^{NE} \leq \tau^{NJ}$ or not at all $\tau^{NE} = 0$. Intuitively, not using the SKG mode for the entire transmission symbol (for example to neutralize the jammer) becomes too costly at high SIR when the jamming interference is relatively low or negligible.

Concerning the uniqueness of the NE, the only case in which the states $(P, 0, \Gamma)$ and $(P, \min\{\tau_{max}, \tau_{th}\}, \Gamma)$ can both be NEs is when the provided utilities are identical, i.e., $\tilde{u}(P, 0, \Gamma) = \tilde{u}(P, \min\{\tau_{max}, \tau_{th}(P)\}, \Gamma)$ in addition to the constraint on the system parameters $1 + \sigma^2\Gamma \geq \sqrt{2}\sigma_H^2 P$ (see Appendix VI-B). However, we argue that such an equality condition on the system parameters can only happen in very special cases, otherwise stated, with zero probability (on a continuous sample space).

Furthermore, whenever player L chooses a strategy of the form $(P, \tau_{th}(P))$ at the NE, the jammer becomes indifferent between all their possible transmit powers in $[0, \Gamma]$ (as per Remark 2). Hence, in such cases, the strategy profile $(P, \tau_{th}(P), \Gamma)$ may not be the unique NE.

Theorem 2: If the legitimate nodes' NE strategy in Theorem 1 is such that $\tau^{NE} = \tau_{th}(P)$, the game $\tilde{\mathcal{G}}$ may have other solutions of the form $(P, \tau_{th}(P), \gamma^{NE})$ with $\gamma^{NE} \in (0, \Gamma)$. More precisely, any strategy of the form $(P, \tau_{th}(P), \gamma^{NE})$ with $\gamma^{NE} \in (0, \Gamma)$ meeting the additional condition $\arg\max_{\tau \in [0,1]} \tilde{u}(P, \tau, \gamma^{NE}) = \tau_{th}(P)$ is also a NE of the game. All such NEs provide identical utility to $\tilde{u}(P, \tau_{th}(P), \Gamma)$.

The proof and the detailed system conditions under which the game may have other NEs of the type $(P, \tau_{th}(P), \gamma^{NE})$ with $\gamma^{NE} \in (0, \Gamma)$ aside from $(P, \tau_{th}(P), \Gamma)$ is provided in Appendix VI-B. These NEs may exist with non-zero proba-bility since the additional condition depends on the variable $\gamma^{NE} \in (0, \Gamma)$ and not only on the system para-meters, as opposed to the condition entailing that $(P, 0, \Gamma)$ and $(P, \min\{\tau_{max}, \tau_{th}\}, \Gamma)$ are both NEs. It suffices that $\arg\max_{\tau \in [0,1]} \tilde{u}(P, \tau, \gamma^{NE}) = \tau_{th}(P)$ holds for a single value of $\gamma^{NE} \in (0, \Gamma)$ to entail the existence of such NEs.

Apart from providing a complete NE analysis, the existence of the NEs in Theorem 2 is not very relevant in practice. First, whenever they exist, the utility at such NEs is identical to the utility of the NE profile: $(P, \tau_{th}(P), \Gamma)$ in Theorem 1. Second, given Remark 2, the jammer can be assumed to restrict their strategy space from $[0, \Gamma]$ to the discrete choices $\{0, \Gamma\}$ with no loss of optimality. Assuming $\tilde{\mathcal{A}}_J = \{0, \Gamma\}$, the resulting game $\tilde{\mathcal{G}}$ has a unique pure-strategy NE (almost surely) which is given in Theorem 1.

As a last result, it turns out that neutralizing the jammer (NJ) in Proposition 1 incurs a non-trivial cost and the obtained utility is lower or equal to the NE utility.

Proposition 2: The SKG utility obtained when neutralizing the jammer (NJ) can never be greater that the utility at the NE. Both utilities are equal, if and only if $\tau^{NE} = \tau_{th}(P)$.

Proof: Since $(P, \tau^{NE}) = \arg\max_{p,\tau} \tilde{u}(p, \tau, \Gamma)$, from the NE's best-response property, we have that $\tilde{u}(p^{NJ}, \tau^{NJ}, \Gamma) \leq \tilde{u}(P, \tau^{NE}, \Gamma)$. From Remark 2, we have that $\tilde{u}(p^{NJ}, \tau^{NJ}, \Gamma) = \tilde{u}(p^{NJ}, \tau^{NJ}, 0)$ (the jammer is indifferent between all its choices) and we obtain that $\tilde{u}(P, \tau^{NE}, \Gamma) \geq \tilde{u}(p^{NJ}, \tau^{NJ}, 0)$. Intuitively, when searching for the NJ state in Proposition 1 the additional condition that the jammer has to be neutralized (i.e., $p = p_{th}(\tau)$) restricts the feasible set of all pairs (p, τ) which results in an optimality loss compared to the NE. Notice that $\max_\tau \tilde{u}(p_{th}(\tau), \tau, 0) \equiv \max_\tau \tilde{u}(p_{th}(\tau), \tau, \Gamma)$. This further implies that, if $\tau^{NE} = \tau_{th}(P)$, the aforementioned restriction is optimal and $(p^{NJ}, \tau^{NJ}) = (P, \tau^{NE})$ which proves the direct implication of the second claim. The hypothesis of the reverse implication: $\tilde{u}(p^{NJ}, \tau^{NJ}, 0) = \tilde{u}(P, \tau^{NE}, \Gamma)$ implies that $\tau^{NE} = \tau_{th}(P)$ and, thus, $\tilde{u}(p^{NJ}, \tau^{NJ}, 0) = \tilde{u}(P, \tau^{NE}, 0)$. From Appendix VI-A, the function $\tilde{u}(p_{th}(\tau), \tau, 0)$ has a unique maximizer w.r.t $\tau \in [0, \tau_{th}(P)]$ given by τ^{NJ} which results in that $(p^{NJ}, \tau^{NJ}) = (P, \tau^{NE})$. ∎

C. Stackelberg Equilibrium

After investigating the NE solution of the strategic interaction in which the legitimate nodes and the jammer choose their optimal strategies simultaneously, a natural rising issue is whether the solution of the game changes assuming a hierarchy in the players' choices [20], [22], [32]. To tackle this issue, we study the SE and compare it to the NE and the jammer neutralization (NJ) states in Sec. III-B and III-A, respectively. We assume that the leader of the game L is playing first by choosing their best action (p^{SE}, τ^{SE}) while anticipating the response of player J. The follower, player J, observes the choice of the leader and reacts optimally (or best-responds) by choosing γ^{SE}.

To be specific, for an arbitrary choice of player L (p, τ), the best-response of the jammer is defined as:

$$\gamma^{BR}(p, \tau) = \arg \min_{\gamma \in [0,\Gamma]} \tilde{u}(p, \tau, \gamma). \tag{17}$$

The leader, anticipating the jammer's reaction described above, can choose their optimal strategy as follows

$$(p^{SE}, \tau^{SE}) = \arg \max_{p,\tau} \tilde{u}(p, \tau, \gamma^{BR}(p, \tau)). \tag{18}$$

The optimal strategy of the jammer is the best response $\gamma^{SE} = \gamma^{BR}(p^{SE}, \tau^{SE})$ given the optimal leader's strategy above. The solution is described in the next Theorem.

Theorem 3: Assuming the hierarchy described above, if $\tau^{NE} < \tau_{th}(P)$ where τ^{NE} is given in Theorem 1, the SE of the game $\tilde{\mathcal{G}}$ is unique (almost surely) and identical to the NE (P, τ^{NE}, Γ). Otherwise, if $\tau^{NE} = \tau_{th}(P)$, both the NJ state in Proposition 1 and the NE (P, τ^{NE}, Γ) are SE solutions providing identical SKG utility.

The proof is included in Appendix VI-C. Notice that in all possible cases $\tau^{NE} \leq \tau_{th}(P)$ (see Theorem 1). The above result shows that neutralizing the jammer is a rational solution when the strategic decisions are not taken simultaneously and the legitimate nodes play first. However, since the NJ state cannot provide a strictly better utility than the NE state (see Proposition 2), the hierarchical play does not bring an actual benefit to player L when compared with the NE.

Finally, we note that as opposed to the NE, the SE requires the leader to be able to anticipate precisely the response of the follower. For this reason, the leader cannot actually choose a strategy such that $p = p_{th}(\tau)$ which renders the follower indifferent between all its actions $\gamma \in [0, \Gamma]$ (and may choose any jamming power in an unpredictable way). A simple way to overcome this issue is for the leader to transmit at $p = p_{th}(\tau) - \varepsilon$ whenever it wants to silence the jammer (at the NJ), and to transmit at $p = p_{th}(\tau) + \varepsilon$ whenever it wants the jammer to transmit at full power (at the NE), with $\varepsilon > 0$ and $\varepsilon \ll 1$ chosen arbitrarily small, with little or no practical impact. Furthermore, this also the excludes other SE solutions (e.g., the NEs in Theorem 2 cannot be SEs).

IV. Channel Hopping vs. Power Spreading in BF AWGN Channels

If the legitimate nodes do not have EH capabilities, we investigate yet another way to defend against jamming by assuming that the legitimate nodes can employ channel hopping or power spreading strategies over multiple orthogonal subcarriers. For this, we generalize the system model (6) and (7) to an N-BF AWGN channel. Alice's and Bob's observations on the i-th subcarrier – denoted by $\hat{Y}_{A,i}$ and $\hat{Y}_{B,i}$ respectively – are expressed as:

$$\hat{Y}_{A,i} = \sqrt{p_i} H_i + \sqrt{\gamma_i} G_{A,i} + Z_{A,i}, \tag{19}$$

$$\hat{Y}_{B,i} = \sqrt{p_i} H_i + \sqrt{\gamma_i} G_{B,i} + Z_{B,i}, \tag{20}$$

where the fading coefficient in the link between Alice and Bob on the i-th subcarrier is denoted by H_i, in the link between Eve and Alice by $G_{A,i}$ and in the link between Eve and Bob by $G_{B,i}$. We assume that the fading coefficients are i.i.d. Gaussian random variables with $H_i \sim \mathcal{N}(0, \sigma_H^2)$, $G_{A,i} \sim \mathcal{N}(0, \sigma_A^2)$ and $G_{B,i} \sim \mathcal{N}(0, \sigma_B^2)$. Notice that the fading coefficients are assumed to have the same statistics. This assumption is justified, since, broadly speaking, narrowband fading depends on the bandwidth (which is the same for all subcarriers) and not on the central frequency (unlike wideband fading or large scale fading) [33]. Furthermore, the noise variables $Z_{A,i}$ and $Z_{B,i}$ are assumed to be i.i.d. Gaussian zero mean unit variance random variables. Finally, Alice and Bob exchange constant probe signals [8] with power p_i and that Eve transmits constant jamming signals [14] with power γ_i on the i-th subcarrier so that the following average power constraints are satisfied[2] [16], [18]:

$$\frac{1}{N}\sum_{i=1}^{N} p_i \leq P, \quad \frac{1}{N}\sum_{i=1}^{N} \gamma_i \leq \Gamma. \tag{21}$$

Given the above model, an easy calculation reveals that the SKG capacity over the i-th subcarrier can be expressed as a function of p_i and γ_i as:

$$\begin{aligned} C(p_i, \gamma_i) &= I(\hat{Y}_{A,i}; \hat{Y}_{B,i}) \\ &= \frac{1}{2}\log_2\left(1 + \frac{\sigma_H^2 p_i}{N_{A,i} + N_{B,i} + \frac{N_{A,i}N_{B,i}}{\sigma_H^2 p_i}}\right), \end{aligned}$$

with

$$N_{A,i} = 1 + \sigma_A^2 \gamma_i, \quad N_{B,i} = 1 + \sigma_B^2 \gamma_i. \tag{22}$$

In order to evaluate the overall SKG capacity, we formalize the channel hopping vs. power spreading techniques similarly to [16] and [18]. When channel hopping is employed, all of the available power is used to transmit on a *single* randomly chosen subcarrier i. Therefore, when the legitimate nodes employ channel hopping on subcarrier i, then $p_i = NP$ and $p_k = 0$ for $k \neq i$, while when the jammer hops on subcarrier i then $\gamma_i = N\Gamma$ and $\gamma_k = 0, k \neq i$. On the other hand, when power spreading is used, the available power is equally distributed across all subcarriers so that $p_i = P$ and $\gamma_i = \Gamma$ $\forall i \leq N$.

When transmitting over the entire spectrum, the choice of the uniform power allocation is motivated by the fact that the nodes do not know their actual channel gains and that their statistics are identical across all frequency carriers. Moreover, assuming that player L transmits with uniform power allocation and from the convexity of the SKG function in (22) w.r.t. γ_i, it turns out that the uniform power allocation for the jammer is optimal and minimizes the overall SKG utility. More general power allocation policies can be considered in future investigations.

From an implementation point of view for the proposed channel hopping and power spreading strategies, we consider that an OFDM transmitter with a standard inverse fast Fourier transform (IFFT) block is employed. In channel hopping mode, all but a randomly chosen IFFT input are set to zero. No coordination regarding the chosen channel hopping or spreading options is required between transmitting and receiving terminals. This is possible if wideband reception is employed by all parties, allowing transmitting terminals to independently choose their strategies without coordination with the receiving terminals. Such a wideband reception of the N orthogonal subcarriers can be efficiently implemented using a standard FFT based OFDM receiver.

Using this framework in the following, for Alice and Bob the probability of channel hopping on subcarrier i is

[2]Using constant probe signals preserves the Gaussianity of the inputs $\sqrt{p_i}H_i$, $\sqrt{\gamma_i}G_{A,i}$ and $\sqrt{\gamma_i}G_{B,i}$, which is optimal for the legitimate nodes and the jammer in our AWGN setting.

denoted by α_i $\forall i \leq N$, while α_{N+1} denotes the probability of spreading the available power uniformly over the whole spectrum. Similarly, we define β_i for the jammer. Since $\alpha = [\alpha_1, \ldots, \alpha_{N+1}]$ and $\beta = [\beta_1, \ldots, \beta_{N+1}]$ are discrete probability distributions, we have the constraints $\alpha_j \geq 0$, $\forall j$, $\sum_{i=1}^{N+1} \alpha_i = 1$, $\beta_j \geq 0$, $\forall j$, and $\sum_{i=1}^{N+1} \beta_i = 1$.

Given the above, the SKG capacity over the N orthogonal subcarriers is given by:

$$\hat{u}(\alpha, \beta) = \frac{1}{N}\Bigg\{ \sum_{i=1}^{N} \{\alpha_i(1-\beta_i-\beta_{N+1})C(NP,0) + \alpha_i\beta_i C(NP, N\Gamma) + \alpha_i\beta_{N+1}C(NP,\Gamma) + \alpha_{N+1}\beta_i[(N-1)C(P,0)+C(P,N\Gamma)]\} + \alpha_{N+1}\beta_{N+1}NC(P,\Gamma)\Bigg\}, \quad (23)$$

where the normalization $\frac{1}{N}$ accounts for measuring the SKG capacity in bits/s/Hz. In (23), the first term corresponds to the case in which Alice (resp. Bob) hops on subcarrier i and the jammer hops on a different subcarrier; the second term to the case in which Alice (resp. Bob) and the jammer both hop on subcarrier i; the third term to the case in which Alice (resp. Bob) hops on subcarrier i and the jammer spreads; the fourth term to the case in which the Alice (resp. Bob) spreads and the jammer hops on subcarrier i. Finally, the last term corresponds to the case in which they both spread their power.

A. Game Formulation and Nash Equilibria

We model the competitive interaction between player L and J as the following zero-sum game $\hat{\mathcal{G}} = \{\hat{\mathcal{A}}_L, \hat{\mathcal{A}}_J, \hat{u}(\alpha,\beta)\}$, where the payoff $\hat{u}(\alpha,\beta)$ is given in (23). The action sets of the players are the probabilities of channel hopping and power spreading:

$$\hat{\mathcal{A}}_L = \left\{ \alpha \in [0,1]^{N+1} \left| \sum_{i=1}^{N+1} \alpha_i = 1 \right. \right\},$$
$$\hat{\mathcal{A}}_J = \left\{ \beta \in [0,1]^{N+1} \left| \sum_{i=1}^{N+1} \beta_i = 1 \right. \right\}.$$

As we have argued in the previous section, the natural solution in such a strategic interaction without cooperation among the opponents is the NE.

To derive the game's NE, let us introduce a finite discrete game $\hat{\mathcal{G}}^D = \{\hat{\mathcal{E}}_L, \hat{\mathcal{E}}_J, \hat{u}(\alpha,\beta)\}$ with action sets defined as $\hat{\mathcal{E}}_L \equiv \hat{\mathcal{E}}_J = \{e_1, \ldots, e_N, e_{(N+1)}\}$, where $e_i \in \{0,1\}^{N+1}$ is the canonical vector containing 1 on the i-th position and 0 otherwise. The i-th action e_i represents channel hopping on subcarrier i for all $i \leq N$ and e_{N+1} represents spreading the power across the spectrum. Such finite discrete games always have at least one NE in mixed strategy (α^*, β^*) [32, Sec. 1.3.1]. We observe that our game $\hat{\mathcal{G}}$ represents the mixed strategy extension of $\hat{\mathcal{G}}^D$ and thus $\hat{\mathcal{G}}$ has at least one NE.

Corollary 1 [32, Th. 1.1]: Game $\hat{\mathcal{G}}$ has at least one NE.

To compute the NEs, one possibility is to use the Minimax Theorem of von Neumann and Morgenstern [34] which allows us to compute mixed NEs of any two-player zero-sum game via linear programming (i.e., by solving two dual linear optimization problems). In our case, we show that the NEs can be characterized in an analytical closed-form manner without the need of solving any optimization problem. To this aim, an alternative characterization of the NE (see Definition 1.2 in [32, Sec.1.2.1]) is used:

Definition 1: A strategy profile $(\alpha^*, \beta^*) \in \hat{\mathcal{A}}_L \times \hat{\mathcal{A}}_J$ is a NE of the game $\hat{\mathcal{G}}$ if the following hold:

i) *both players are indifferent among the pure actions that are played with positive probability at the NE*

$$\hat{u}(\alpha^*, e_i) = \hat{u}(\alpha^*, e_k), \quad \forall i,k, \in \mathcal{I}_J,$$
$$\hat{u}(e_i, \beta^*) = \hat{u}(e_k, \beta^*), \quad \forall i,k, \in \mathcal{I}_L,$$

ii) *the pure actions that result in strictly smaller payoffs are played with zero probability at the NE*

$$\text{if } \hat{u}(\alpha^*, e_i) < \hat{u}(\alpha^*, e_k), \quad i \in \mathcal{I}_J, \quad \text{then } k \in \mathcal{N}_J,$$
$$\text{if } \hat{u}(e_i, \beta^*) > \hat{u}(e_k, \beta^*), \quad i \in \mathcal{I}_L, \quad \text{then } k \in \mathcal{N}_L,$$

where the sets $\mathcal{N}_L, \mathcal{I}_L \subseteq \{1, \ldots, N+1\}$ denote, respectively, the indices of the pure actions that are not played at the NE and those that are played at the NE by player L: $\mathcal{N}_L = \{i \,|\alpha_i^* = 0\}$, $\mathcal{I}_L = \{1, \ldots, N+1\} \setminus \mathcal{N}_L$; similarly, the sets $\mathcal{N}_J, \mathcal{I}_J \subseteq \{1, \ldots, N+1\}$ denote, respectively, the set of indices of the pure actions that are not used or are used by player J at the NE: $\mathcal{N}_J = \{i \,|\beta_i^* = 0\}$, and $\mathcal{I}_J = \{1, \ldots, N+1\} \setminus \mathcal{N}_J$.

At a first glance, Definition 1 provides a simple way to compute the NEs of the game $\hat{\mathcal{G}}$ by solving a system of linear equations and checking some conditions. Still, in order to use Definition 1, one would have to know in advance the faces of the simplex $\hat{\mathcal{A}}_L \times \hat{\mathcal{A}}_J$ on which the NEs lie, i.e., one would have to know $\mathcal{I}_L$, $\mathcal{I}_J$ for all NEs. An exhaustive search has an exponential complexity (the $N+1$-simplex has $2^{N+1}-1$ faces). Nevertheless, the NEs of our game $\hat{\mathcal{G}}$ have a special structure which allows us to exploit Definition 1 and fully characterize the set of NEs in a simple manner.

To characterize the set of NEs as a function of the system's parameters we begin by examining the matrix structure of the discrete game $\hat{\mathcal{G}}^D$ given in Table I. We notice that there is a symmetry between the channel hopping strategies. In particular, the payoff does not depend on the particular index of the chosen subcarrier but only on whether both players hop on the same subcarrier or not. This symmetry allows us to show that the NE of the game $\hat{\mathcal{G}}$ have a particular structure specified in the following propositions.

Proposition 3: At the NE (α^*, β^*), a player uses either all channel hopping actions with non-zero probability or none of them: either $\alpha_i^* = 0$, $\forall i \leq N$ or $\alpha_i^* \neq 0$, $\forall i \leq N$, and similarly, either $\beta_i^* = 0$, $\forall i \leq N$ or $\beta_i^* \neq 0$, $\forall i \leq N$.

Proposition 4: If both players employ channel hopping with non-zero probability at the NE, i.e., $\alpha_i^* > 0$ and $\beta_i^* > 0$ $\forall i \leq N$, then the players will hop uniformly across all channels and the NE will have the following structure: $\alpha^* = (a, \ldots, a, (1-Na))$, $\beta^* = (b, \ldots, b, (1-Nb))$ for some $0 \leq a \leq 1/N$, $0 \leq b \leq 1/N$.

Propositions 3 and 4 are proven in Appendices VI-D and VI-E. These results shape the special structure of the NEs of $\hat{\mathcal{G}}$, which, alongside Definition 1 and the strict convexity of $C(p,\gamma)$ w.r.t. γ, allows us to fully

TABLE I
TWO PLAYER ZERO-SUM DESCRIPTION OF $\hat{\mathcal{G}}_d$

	$e_i,\ i \leq N$	$e_k,\ k \leq N, k \neq i$	e_{N+1}
$e_i,\ i \leq N$	$\frac{1}{N}C(NP, N\Gamma)$	$\frac{1}{N}C(NP, 0)$	$\frac{1}{N}C(NP, \Gamma)$
$e_k,\ k \leq N, k \neq i$	$\frac{1}{N}C(NP, 0)$	$\frac{1}{N}C(NP, N\Gamma)$	$\frac{1}{N}C(NP, \Gamma)$
e_{N+1}	$\frac{N-1}{N}C(P, 0) + \frac{1}{N}C(P, N\Gamma)$	$\frac{N-1}{N}C(P, 0) + \frac{1}{N}C(P, N\Gamma)$	$C(P, \Gamma)$

characterize the set of NEs in a very simple and explicit manner as function of the system parameters.

Theorem 4: The set of NEs of the game $\hat{\mathcal{G}}$ is characterized as follows:

1. *If $C(NP, \Gamma) < NC(P, \Gamma)$, then the game has a unique pure-strategy NE: both players spread their powers, $\alpha^* = \beta^* = e_{N+1}$.*
2. *If $C(NP, \Gamma) > NC(P, \Gamma)$, then player L hops and player J spreads at the NE: $\alpha^* = (\alpha_1, \ldots, \alpha_N, 0)$ and $\beta^* = e_{N+1}$. The NE strategies of player L are given by the (infinite number of) solutions to the following system of linear inequalities:*

$$\begin{cases} 0 \leq \alpha_i \leq 1, \quad \forall i \leq N, \\ \sum_{j=1}^{N} \alpha_j = 1, \\ \alpha_i < \dfrac{C(NP, 0) - C(NP, \Gamma)}{C(NP, 0) - C(NP, N\Gamma)}, \quad \forall i \leq N. \end{cases}$$

   In particular, the uniform probability over the channels is one of the NE solutions: $\alpha^* = (1/N, \ldots, 1/N, 0)$. All NEs are equivalent in terms of achieved utility.
3. *If $C(NP, \Gamma) = NC(P, \Gamma)$, player L employs all their actions and player J spreads at the NE: $\alpha^* = (\alpha_1, \ldots, \alpha_N, \alpha_{N+1})$ and $\beta^* = e_{N+1}$. The NE strategies of player L are the (infinite number of) solutions to the following linear system of inequalities:*

$$\begin{cases} \alpha_i \geq 0, \quad \forall i \leq N, \\ \sum_{j=1}^{N} \alpha_j = 1, \\ \alpha_i[C(NP, N\Gamma) - C(NP, 0)] + \alpha_{N+1}[(N-1)C(P, 0) \\ +C(P, N\Gamma) - C(NP, 0) + C(NP, \Gamma) - NC(P, \Gamma)] \\ > C(NP, \Gamma) - C(NP, 0), \quad \forall i \leq N. \end{cases}$$

   In this case, both players spreading (case 1) is an NE. Also, player J spreading and player L hopping strategies (case 2) are all NEs. All NEs are equivalent in terms of achieved utility.

The proof is provided in Appendix VI-F. We remark that the NE can be unique and in pure strategies if $C(NP, \Gamma) < NC(P, \Gamma)$ and the outcome of the game provides a utility equal to $\hat{u}(\alpha^*, \beta^*) = C(P, \Gamma)$. On the contrary, if $C(NP, \Gamma) \geq NC(P, \Gamma)$, there are an infinite number of NEs which are generally in mixed strategies. All these NEs are equivalent in terms of achieved utility, which equals $\hat{u}(\alpha^*, \beta^*) = \frac{1}{N}C(NP, \Gamma)$. Hence, the outcome of the game can be predicted without the need for implementing iterative or learning procedures.

Theorem 4 also shows that the optimal strategy for the jammer is always spreading their power across the entire spectrum. Intuitively, if the jammer were to use channel hopping, player L would exploit this fact and would also hop; this scenario is unfavorable for the jammer as the probability that both players hop on the same subcarrier equals $\frac{1}{N^2}$ (due to Proposition 3, when both players hop at the NE, they use uniform probabilities). Thus, the jammer's payoff from hopping cannot exceed that gained from spreading, assuming that the legitimate nodes play their optimal strategy. On the contrary, for player L the best strategy can be either channel hopping or power spreading depending on which option provides higher utility against a spreading jammer.

B. Stackelberg Equilibrium

In Sec. III-C, we have shown that the hierarchy of play among the adversaries does not bring an advantage to the legitimate nodes assuming they have EH capabilities. Here, we investigate whether this remains true in N-BF AWGN systems in which the players choose between channel hopping and power spreading strategies. The leader, player L, is assumed to play first and to choose α^{SE} anticipating the jammer's response. The follower, player J, observes α^{SE} and best-responds by choosing β^{SE}.

More precisely, the best-response of the jammer for an arbitrary choice of α is defined as: $\beta^{BR}(\alpha) = \arg\min_{\beta} \hat{u}(\alpha, \beta)$. Thus, the leader chooses their optimal strategy as follows

$$\alpha^{SE} = \arg\max_{\alpha} \ \hat{u}(\alpha, \beta^{BR}(\alpha)) \tag{24}$$

and the resulting best-response or SE strategy of the jammer is $\beta^{SE} = \beta^{BR}(\alpha^{SE})$.

To characterize the SE in closed-form, we use a similar approach as for the NE: we show first that the leader's strategy at the SE has a special form described below. Then, we exploit this structure to provide the SE solution.

Proposition 5: At the SE, the legitimate player uses either all hopping strategies with uniform probability or none of them, i.e., $\alpha^{SE} = (a, \ldots, a, 1 - Na)$ for some $a \in [0, 1/N]$.

The proof is provided in Appendix VI-G. The above structure of α^{SE} allows us to analyze the optimal response of the jammer β^{SE} and to prove that, in all cases, the jammer's best strategy is to spread: $\beta^{SE} = (0, \ldots, 0, 1)$. On the other hand, depending on the channel parameters, the leader will either channel hop or spread their power, identically to the NE.

Theorem 5: The set of SEs of the game $\hat{\mathcal{G}}$ is identical to the set of NEs.

The proof is provided in Appendix VI-H. Therefore, the legitimate nodes do not gain in utility by choosing first their strategy as opposed to the NE where both players choose their strategies simultaneously.

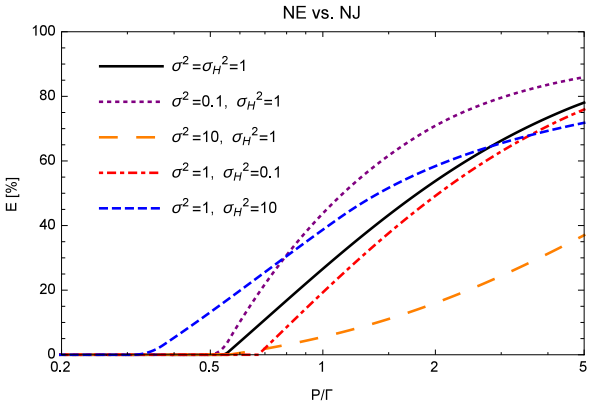


Fig. 2. Relative utility gain at the NE vs. NJ $E = (C^{NE} - C^{NJ})/C^{NE}$ as a function of $P/\Gamma \geq 0$ for $\zeta = 0.7$.

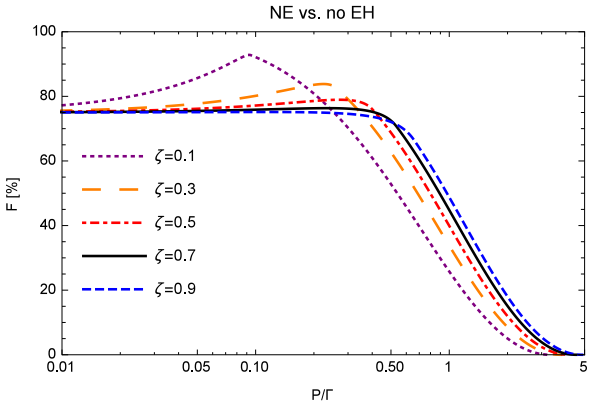


Fig. 3. Relative utility gain at the NE vs. no EH: $F = (C^{NE} - C^{noEH})/C^{NE}$ as a function of $P/\Gamma \geq 0$.

V. Numerical Illustrations and Discussion

In this Section, several representative illustrations are chosen allowing the deduction of generic conclusions that carry over most setups. The benchmark setting is chosen as follows: unit jamming power $\Gamma = 1$, unit variance Rayleigh channel coefficients $\sigma_A^2 = \sigma_B^2 = \sigma^2 = \sigma_H^2 = 1$.

A. EH at the Legitimate Nodes

We start by evaluating the SKG capacity at the NJ in Proposition 1 and NE in Theorem 1 as functions of the system parameters for a harvesting efficiency $\zeta = 0.7$. In Fig. 2, the relative gain in utility obtained at the NE ($C^{NE} = \tilde{u}(P, \tau^{NE}, \Gamma)$) compared with the NJ ($C^{NJ} = \tilde{u}(p^{NJ}, \tau^{NJ}, 0)$) defined by $E \triangleq \frac{C^{NE}-C^{NJ}}{C^{NE}}$, is depicted as a function of the signal to interference ratio (SIR) P/Γ for different values of σ^2 and σ_H^2. In the investigated settings, the NJ strategy never outperforms the NE in terms of utility, which is consistent with Proposition 2. When the SIR P/Γ is relatively low, both the NE and the NJ provide identical utilities. In this case, $p^{NJ} = P$ and $\tau^{NJ} = \tau^{NE} = \tau_{th}(P)$, the jammer is indifferent between $\{0, \Gamma\}$ and both states are SE solutions. With increasing SIR P/Γ, it is no longer optimal for the legitimate nodes to harvest energy for a fraction of time $\tau_{th}(P)$ in order to neutralize the jammer. Instead, by limiting the duration of EH to $\tau^{NE} = \tau_{max} < \tau_{th}(P)$ the SKG capacity increases in spite of the full power jamming $\gamma = \Gamma$ and only the NE is also a SE solution. Moreover, as the SIR increases, e.g., for $P/\Gamma \gg 1$, the legitimate nodes should not harvest energy at all as the jammer's interference is relatively negligible.

Notice that Fig. 2 also illustrates the SE solution described in Theorem 3. Indeed, at low SIR, when both NE and NJ provide equal SKG capacity, they are both SE solutions. At high SIR, the SE is unique and identical to the NE.

Subsequently, we evaluate the impact of the EH capability on the SKG capacity at the NE. The relative gain in utility obtained at the NE compared with the case in which there is no EH capability $C^{NoEH} = \tilde{u}(P, 0, \Gamma) = C(P, \Gamma)$, defined as $F \triangleq \frac{C^{NE}-C^{NoEH}}{C^{NE}}$, is depicted as a function of P/Γ in Fig. 3. The benchmark setup is considered and the different curves correspond to harvesting efficiencies $\zeta \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$. At low SIR, F decreases with the harvesting efficiency ζ. Although counter-intuitive, this is explained by the fact that, at the NE, the harvesting period $\tau^{NE} = \tau_{th}(P) = P/\zeta$ decreases in ζ in this regime. The peaks represent the transition points to the second regime, in which harvesting energy at the threshold is no longer optimal and $\tau^{NE} = \tau_{max} < \tau_{th}(P)$. Also, as the SIR increases, the curves progressively switch ranks and F becomes increasing in ζ as expected. For $P/\Gamma = 1$ and $\zeta = 0.3$ the gain in using EH is around 30 % while it increases to 40 % for $\zeta = 0.7$. At low SIR P/Γ the gain observed can reach 90 %, while at the high SIR it is negligible: in the third regime $\tau^{NE} = 0$, as harvesting energy becomes time-consuming and inefficient in terms of SKG capacity.

Finally, the relative utility F defined above is depicted in Fig. 4 for $\zeta = 0.7$ and various channel parameters. For low SIR P/Γ, there is a significant gain in utility when employing EH. This gain becomes significantly large at very low SIR, exceeding 97.5 % when the legitimate nodes experience poor channel conditions as opposed to the jammer. When both parties experience similar channel conditions the gain is in the range of 60 % in the medium SIR. Overall, the numerical results demonstrate that using EH as a counter-jamming technique is of particular interest in the low and medium SIR regimes but, as expected, does not increase the utility in the high SIR. The peaks represent here as well the transition from the $\tau^{NE} = \tau_{th}(P)$ regime (at low SIR) to the second regime in which $\tau^{NE} = \tau_{max} < \tau_{th}(P)$.

B. Channel Hopping vs. Power Spreading

First, we analyze the NE as function of N and the ratio P/Γ for the benchmark scenario in Fig. 5. There exist two regions delimited by the curve $C(NP, \Gamma) = NC(P, \Gamma)$: a region in which the NE is unique and both players spread their power, and a region in which the jammer spreads their power and the legitimate nodes employ channel hopping.

Player L hops at the NE below the curve, when the SIR P/Γ is relatively small. This is intuitive since, in the low

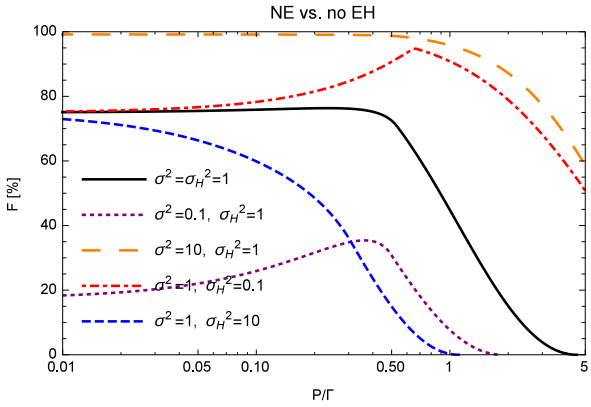


Fig. 4. Relative utility gain at the NE vs. no EH: $F = (C^{NE} - C^{noEH})/C^{NE}$ as a function of $P/\Gamma \geq 0$ for $\zeta = 0.7$ and different channel parameters.

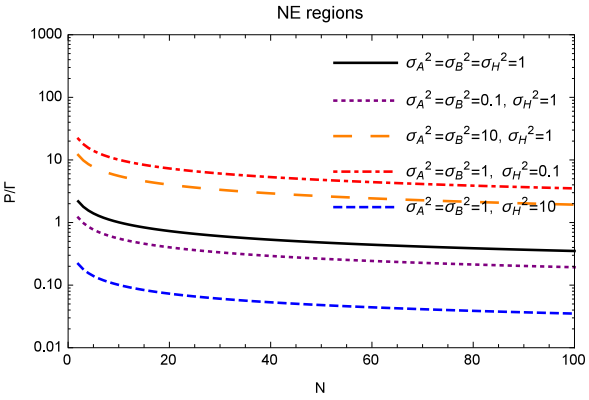


Fig. 6. NE regions as a function of $P/\Gamma \geq 0$ and $N \geq 2$ for $\Gamma = 1$ and different channel parameters.

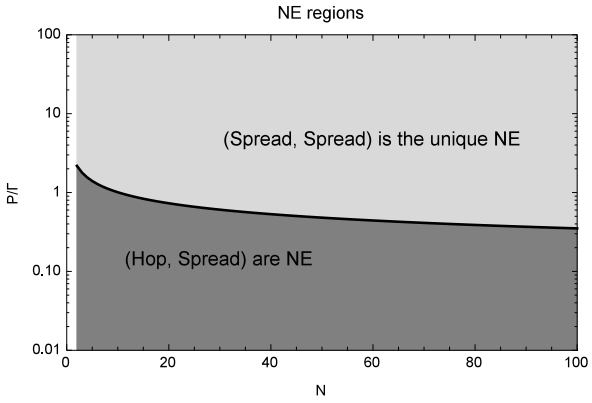


Fig. 5. NE regions as a function of $P/\Gamma \geq 0$ and $N \geq 2$ for $\Gamma = \sigma_A^2 = \sigma_B^2 = \sigma_H^2 = 1$.

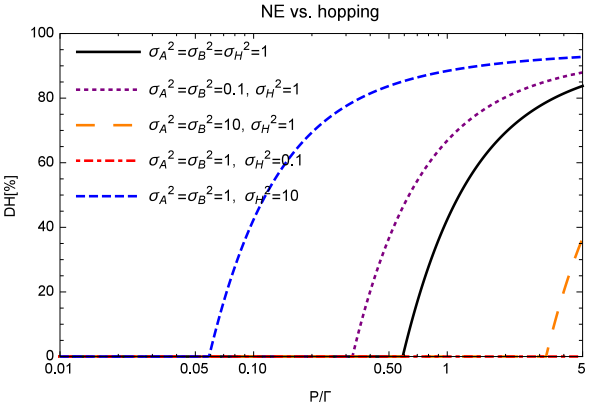


Fig. 7. Relative utility gain between the NE vs. always hopping: $D_H = (u^{NE} - u^{Hop,Spread})/u^{NE}$ as a function of P/Γ for $N = 32$, $\Gamma = 1$ and different channel parameters.

transmit power regime, the legitimate nodes should avoid as much jamming interference as possible by transmitting on a single subcarrier, which also means that their available power is concentrated on a single channel.

In Fig. 6, the NE regions are illustrated for different channel parameters. When σ_H^2 increases, the region in which player L should employ channel hopping at the NE shrinks down while when σ_A^2, σ_B^2 increase, this region expands.

In Fig. 7, the relative gain obtained by player L when employing the NE strategy as opposed to a naive hopping strategy is depicted. The relative utility gain $D_H = (u^{NE} - u^{Hop,Spread})/u^{NE}$, where $u^{Hop,Spread} = 1/NC(NP, \Gamma)$ is relatively large (up to 80%) in the high SIR regime, in which case the optimal strategy for player L is to use the entire spectrum in spite of the jammer's interference.

Finally, in Fig. 8, the relative utility gain when using the NE strategy over N subcarriers as opposed to a single subcarrier ($u^{single} = C(P, \Gamma)$) is investigated for $\Gamma = \sigma_H^2 = \sigma_A^2 = \sigma_B^2 = 1$ as a function of P/Γ for $N \in \{2, 4, 8, 16, 32, 64\}$. At low SIR, when the channel hopping strategy is optimal for the legitimate nodes, the higher the number of subcarriers N, the lower the jammer's interference in each subcarrier, and hence, the higher the SKG capacity. At last, in the high SIR regime, when spreading is optimal the SKG utility becomes $C(P, \Gamma)$, which is identical to transmitting over a single channel with powers P and Γ.

Remark that all figures illustrating the NE, in this subsection, also illustrate the SE solution, since the SE is identical to the NE as per Theorem 5.

C. Discussion and Perspectives

We discuss here the differences and similarities between the two approaches: a) EH at the legitimate nodes, and b) employing channel hopping or power spreading techniques.

EH at the legitimate nodes enables them to completely neutralize the jammer. By harvesting the jamming power in a first phase and exploiting it for SKG in a second phase, the jammer's attacks may increase the SKG capacity; in this case, the jammer should not launch the attack, i.e., it is neutralized. However, it is not always optimal for the legitimate nodes to neutralize the jammer. Indeed, using EH can reduce the SKG capacity since, for a non-trivial fraction of time, there is no secret bits generation; when the jammer is neutralized the

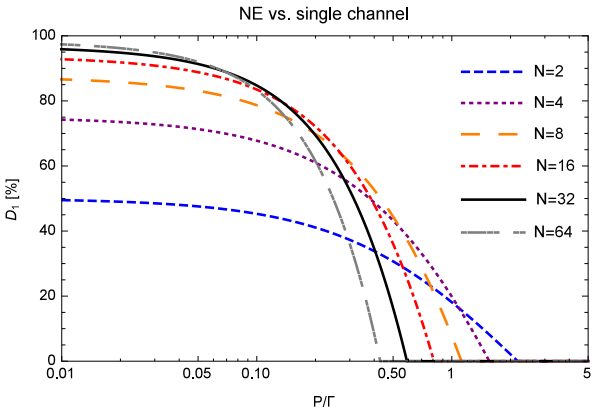


Fig. 8. Relative utility gain between the NE vs. single channel SKG: $D_1 = (u^{NE} - u^{single})/u^{NE}$ as a function of P/Γ for $\Gamma = \sigma_H^2 = \sigma_A^2 = \sigma_B^2 = 1$ and $N \in \{2, 4, 8, 16, 32, 64\}$.

penalty in terms of utility might become too high, depending on the system parameters (e.g., high SIR regime). In such cases, the obtained utility at the NE is strictly higher than the one at the NJ state.

On the other hand, in the case of BF AWGN channels (i.e., in systems with multiple orthogonal subcarriers), the idea is to use channel hopping in a random fashion and avoid most of the jammer's interference as opposed to completely neutralizing it. Since potential jammers cannot predict the subcarrier used by the legitimate nodes, they will always spread their powers over the entire spectrum: the larger the number of subcarriers, the smaller the jammer's interference on each subcarrier. However, channel hopping is not always optimal since only a fraction of the entire spectrum is used for SKG. Depending on the system parameters (high SIR), it can be preferable for the legitimate nodes to spread the available power across the entire spectrum rather than concentrate it on a single subcarrier. In this case, the SKG capacity (measured in bits/s/Hz) is identical to that of single channel with the same average power constraints.

In the critical cases of low and medium SIR regimes ($P/\Gamma < 1$), both approaches turn out to be advantageous in terms of SKG capacity compared to a single channel SKG system without EH capabilities; the gains in SKG capacity depend on the harvesting duration or the number of subcarriers N. On the contrary, in the high SIR regime ($P/\Gamma \gg 1$), the jammer's impact and interference become relatively low or even negligible and the cost of counter-jamming measures might not be justified compared to simply tolerating it. However, the interesting cases are indeed the former ones in which the jamming power is higher or of the same order as the legitimate nodes' transmit powers, in which overcoming the attack becomes critical.

For both approaches it turns out that a hierarchical decision model that in principle could favor the legitimate nodes compared to a simultaneous decision model does not bring an actual benefit. Indeed, the SKG utility obtained at the SE is identical to the SKG utility at the NE (even though the set of SEs is not necessarily identical to the set of NEs as in the EH approach).

Several questions arise for future work. First, an interesting issue would be to study reactive vs. proactive jamming [35] as well as the joint use of EH and multi-carrier transmission against jamming attacks. Second, in the EH case, the study of more realistic models accounting for finite storage capabilities, asymmetries in the legitimate nodes' parameters and EH at the jammer side, are interesting future extensions. Moreover, the study of multi-user and multi-jammer interactions as well as games of incomplete information are challenging open issues.

VI. Conclusions

In this work, the adversarial interaction between a pair of legitimate nodes and a malicious jammer in a wireless secret key generation (SKG) framework was investigated. Two different counter-jamming approaches were proposed and studied. First, energy harvesting at the legitimate nodes, and, second, channel hopping vs. power spreading in block fading AWGN channels. In either approach, a zero-sum game was introduced as the objectives of the two parties involved were opposed. Complete characterizations of the Nash and Stackelberg equilibria in closed-form were provided in both cases. It was demonstrated that either approach may offer significant gains in utility, particularly in the low signal-to-interference ratio regime, in which counteracting the jamming interference becomes crucial. As a result, viable and low complexity alternatives for defending SKG systems may be developed by exploiting either novel transceiver features or available spectral resources.

Appendix

A. Proof of Proposition 1

Let us assume that the legitimate nodes neutralize the jammer by transmitting at power $p \in [0, \min\{p_{th}(\tau), P\}]$. The jammer observes player L's choice and from Remark 2, decides to stay silent. Notice that player L can force the jammer to remain silent by transmitting at $p \in [0, \min\{p_{th}(\tau) - \varepsilon, P\}]$ for an arbitrarily small $\varepsilon > 0$. For simplicity, $\varepsilon \simeq 0$ is assumed in the following.

The remaining question is: how will player L choose $\tau \in [0, 1)$ and $p \in [0, \min\{p_{th}(\tau), P\}]$ to maximize the resulting SKG utility

$$\tilde{u}(p, \tau, 0) = \frac{1-\tau}{2} \log_2 \left(1 + \frac{p\sigma_H^2}{2(1-\tau) + \frac{(1-\tau)^2}{p\sigma_H^2}} \right), \quad (25)$$

while ensuring that the jammer stays silent and cannot decrease the utility by transmitting with non-zero power? Since the feasible set of p depends on τ, we first have to find the maximum of $\tilde{u}(p, \tau, 0)$ w.r.t. p for any fixed τ. The function $\tilde{u}(p, \tau, 0)$ is strictly increasing in p and, hence, the optimal power is given by $\tilde{p}(\tau) = \min\{P, p_{th}(\tau)\}$. Now, we need to maximize $\tilde{u}(\tilde{p}(\tau), \tau, 0)$ w.r.t. $\tau \in [0, 1]$:

$$\tilde{u}(p_{th}(\tau), \tau, 0) = \frac{1-\tau}{2} \log_2 \left(1 + \frac{\zeta\tau\sigma_H^2}{(2 + \frac{1-\tau}{\zeta\tau\sigma_H^2})(1-\tau)} \right).$$

At the extremes $\tau = 0$ and $\tau \to 1$ the utility goes to zero. By investigating its second order derivatives w.r.t. τ, which amounts to the following quadratic equation:

$$(1-\tau)^2 - 2\sigma_H^4 \zeta^2 \tau^2 = 0, \tag{26}$$

it can be shown that $\tilde{u}(p_{th}(\tau), \tau, 0)$ always has an inflexion point in between $(0, 1)$ and starts as convex and then becomes concave. Knowing that the the utility is always positive, we can conclude that $\tilde{u}(p_{th}(\tau), \tau, 0)$ has a unique critical point that is the global maximizer $\tau^* \in (0, 1)$ and which is the solution to $\frac{d\tilde{u}(p_{th}(\tau),\tau,0)}{d\tau} = 0$. This implies that, if $p_{th}(\tau^*) \leq P$, then the optimal solution that neutralizes the jammer is $\tau^{NJ} = \tau^*$ and $p^{NJ} = p_{th}(\tau^*)$. If $p_{th}(\tau^*) > P$ (or equivalently $\tau^* > \tau_{th}(P)$), then the optimal solution that neutralizes the jammer is $p^{NJ} = P$ and $\tau^{NJ} = \tau_{th}(P) = \frac{P}{\zeta}$.

B. Proof of Theorem 1 and Theorem 2

From Remark 1, we know that transmitting at maximum power is a strictly dominant strategy for player L and, hence, $p^{NE} = P$. We first prove that at the NE, player L will not operate in EH mode for longer than the threshold $\tau_{th}(P)$. Let's suppose by absurdum that $\tau^{NE} > \tau_{th}(P)$, then the jammer's best response would be to remain silent $\gamma^{NE} = 0$ (as the energy harvested from the jammer in the EH phase is enough to overcome the interference inflicted by the jammer in the SKG phase). Then, the optimal τ^{NE} maximizing the utility $\tilde{u}(P, \tau, 0)$ (which is decreasing in τ) would be $\tau^{NE} \to \tau_{th}(P)$ obtaining the utility $\tilde{u}^{NE} \to \tilde{u}(P, \tau_{th}(P), 0)$. However, this state cannot be an NE. Indeed, if the jammer stays silent $\gamma^{NE} = 0$, no energy is harvested during τ^{NE} and player L gains in utility by deviating to $\tau = 0$. This will also cause the jammer to deviate to $\gamma = \Gamma$.

The above implies that player L can only choose an EH strategy such that $\tau^{NE} \leq \tau_{th}(P)$ at the NE. This condition is equivalent to $P \geq p_{th}(\tau^{NE})$, which means that the utility is either decreasing or simply a constant in γ (see Remark 2). This further implies that if the jammer uses maximum power $\gamma^{NE} = \Gamma$, then it cannot benefit by deviating unilaterally. Hence, to find the NE of the form (P, τ^{NE}, Γ), we only need to find the optimal value or values of $\tau \in [0, \tau_{th}(P)]$ that maximizes the function $\tilde{u}(P, \tau, \Gamma)$ given by:

$$\tilde{u}(P, \tau, \Gamma) = \frac{1-\tau}{2} \log_2 \left(1 + \frac{\left(\frac{P}{1-\tau} + \kappa\Gamma\right)\sigma_H^2}{2(1+\sigma^2\Gamma) + \frac{(1+\sigma^2\Gamma)^2}{\left(\frac{P}{1-\tau}+\kappa\Gamma\right)\sigma_H^2}} \right),$$

where κ depends on τ: $\kappa(\tau) = \frac{\zeta\tau\sigma^2}{1-\tau}$. At $\tau = 0$, this function is strictly positive $\tilde{u}(P, 0, \Gamma) > 0$ equal to the SKG capacity without EH and, when $\tau \to 1$ the function goes to 0. By investigating the second order derivative of $\tilde{u}(P, \tau, \Gamma)$ w.r.t. τ, which amounts to the analysis of the following quadratic equation

$$(1-\tau)^2(1+\sigma^2\Gamma)^2 - 2\sigma_H^4(P + \sigma^2\zeta\Gamma\tau)^2 = 0, \tag{27}$$

two different cases arise:

- *Case A:* If $1+\sigma^2\Gamma \geq \sqrt{2}\sigma_H^2 P$, this function has a unique inflexion point that lies in $(0, 1)$ and the function starts as convex and then becomes concave. Thus, $\tilde{u}(P, \tau, \Gamma)$ has a critical point that is a local maximum $\tau_{max} \in (0, 1)$, which is a solution of the equation $\frac{d\tilde{u}(P,\tau,\Gamma)}{d\tau} = 0$. Hence, the optimal strategy is given by either the maximal point τ_{max} or by one (or both) of the borders of the interval $[0, \tau_{th}(P)]$ depending on the system parameters:

$$\tau^{NE} = \arg \max_{\tau \in \{0, \min\{\tau_{th}(P), \tau_{max}\}\}} \tilde{u}(P, \tau, \Gamma). \tag{28}$$

- *Case B:* If $1 + \sigma^2\Gamma < \sqrt{2}\sigma_H^2 P$, then the function is always concave (and it does not have an inflexion point) in $(0, 1)$. If the function has a critical point in $(0, 1)$, then this critical point is a maximum point denoted by τ_{max} and $\tau^{NE} = \min\{\tau_{th}(P), \tau_{max}\}$. Otherwise, the function is concave decreasing and $\tau^{NE} = 0$.

Remark that, at least in theory, Case A can lead to the existence of two NEs whenever the additional equality condition is met: $\tilde{u}(P, 0, \Gamma) = \tilde{u}(P, \min\{\tau_{th}(P), \tau_{max}\}, \Gamma)$, i.e., when both borders of the interval $[0, \min\{\tau_{max}, \tau_{th}(P)\}]$ provide equal maximum utility. However, this can happen only in very special cases of the system parameters or with zero probability.

Aside from the zero probability case described above, this profile may not be the unique NE of the game $\tilde{\mathcal{G}}$ as there may exist other NEs such that $\gamma^{NE} \in [0, \Gamma)$. Such cases can only happen if the strategy of player L at the NE equals $(P, \tau_{th}(P))$ or equivalently if $(P, \tau_{th}(P), \Gamma)$ is the above NE. Otherwise, whenever $\tau^{NE} < \tau_{th}(P)$, the utility is strictly decreasing in γ and the only strategy of the jammer at the NE is Γ (case discussed previously).

Now, whenever the legitimate user chooses their strategy $(P, \tau_{th}(P))$, the jammer becomes totally indifferent between all their strategies and, in particular, all jamming powers in $[0, \Gamma)$ provide the same utility (see Remark 2). Hence, in this case, there may be other NEs aside from $(P, \tau_{th}(P), \Gamma)$ that provide identical utilities to $\tilde{u}(P, \tau_{th}(P), \Gamma)$. We can disregard the state $(P, \tau_{th}(P), 0)$ for the same reasons for which the NJ state is not a NE.

In order to find all NE of the form $(P, \tau_{th}(P), \gamma^{NE})$, we need to find all $\gamma^{NE} \in (0, \Gamma)$ such that the legitimate user cannot deviate from $(P, \tau_{th}(P))$ or it will lose in terms of utility. Stated otherwise, all $\gamma^{NE} \in (0, \Gamma)$ such that $\tau_{th}(P) = \arg\max_\tau \tilde{u}(P, \tau, \gamma^{NE})$ provide additional NE profiles of the form $(P, \tau_{th}(P), \gamma^{NE})$.

The analysis of the utility $\tilde{u}(P, \tau, \gamma^{NE})$ as a function of τ is very similar to $\tilde{u}(P, \tau, \Gamma)$ above. There are two cases in function of the system parameters.

- *Case A:* If $1 + \sigma^2\Gamma \geq \sqrt{2}\sigma_H^2 P$, for all $\gamma^{NE} \in \left[\frac{\sqrt{2}\sigma_H^2 P - 1}{\sigma^2}, \Gamma\right)$, the function $\tilde{u}(P, \tau, \gamma^{NE})$ has a unique inflexion point that lies in $(0, 1)$ and starts as convex and then becomes concave. Thus, $\tilde{u}(P, \tau, \gamma^{NE})$ has a critical point that is a local maximum $\tau_{max}(\gamma^{NE}) \in (0, 1)$, which is a solution of the equation $\frac{d\tilde{u}(P,\tau,\gamma^{NE})}{d\tau} = 0$. The additional conditions for the strategy $(P, \tau_{th}(P))$ to be optimal for player L are:

$$\begin{aligned} \tau_{th}(P) &\leq \tau_{max}(\gamma^{NE}) \\ \tilde{u}(P, 0, \gamma^{NE}) &\leq \tilde{u}(P, \tau_{th}(P), \gamma^{NE}) \end{aligned} \tag{29}$$

- *Case B:* If $1+\sigma^2\Gamma < \sqrt{2}\sigma_H^2 P$, for all $\gamma^{NE} \in \left(0, \frac{\sqrt{2}\sigma_H^2 P-1}{\sigma^2}\right)$, the function $\tilde{u}(P,\tau,\gamma^{NE})$ is always concave in $\tau \in (0,1)$. If the function has a critical point in $(0,1)$, then this critical point is a maximum point denoted by $\tau_{max}(\gamma^{NE})$. The additional condition for the strategy $(P,\tau_{th}(P))$ to be optimal is $\tau_{th}(P)=\tau_{max}(\gamma^{NE})$. Otherwise, the function is concave decreasing in τ and $(P,\tau_{th}(P))$ cannot be optimal for player L.

C. Proof of Theorem 3

Let us first find the best-response of the jammer defined in (17). Given the second remark, it is easy to see that:

$$\gamma^{BR}(p,\tau)=\begin{cases}0, & \text{if } p<p_{th}(\tau)\\ \in[0,\Gamma], & \text{if } p=p_{th}(\tau)\\ \Gamma, & \text{if } p>p_{th}(\tau),\end{cases} \tag{30}$$

where $p_{th}(\tau)=\zeta\tau$. Notice that whenever $p=p_{th}(\tau)$ the best response of the jammer can be anything and cannot in fact be predicted by player L. However, the obtained payoff is anticipated by player L as it does not depend on the actual choice of the jammer: $\tilde{u}(p_{th}(\tau),\tau,\gamma)=\tilde{u}(p_{th}(\tau),\tau,0)$, for all γ.

The SE action of the leader, anticipating that the jammer will best respond to their own choice is given by:

$$(p^{SE},\tau^{SE})=\arg\max_{p,\tau}\ \tilde{u}(p,\tau,\gamma^{BR}(p,\tau)) \tag{31}$$

From (30), we see that player L can either neutralize the jammer or allow it to transmit, knowing that the jammer will transmit with full power Γ. The situation that proves to be mostly advantageous to the legitimate player will be chosen.

- *Case A:* Assume the legitimate player neutralizes the jammer by choosing a strategy such that $p\le p_{th}(\tau)$. Player L has to find the best pair (p,τ) that maximizes $\tilde{u}(p,\tau,0)$ knowing that $p\in[0,\min\{P,p_{th}(\tau)\}]$ and that $\tau\in[0,1]$; the solution equals (p^{NJ},γ^{NJ}) in Proposition 1.

- *Case B:* Assume now that the legitimate player does not neutralize the jammer and $p\ge p_{th}(\tau)$. Player L has to find the best pair (p,τ) that maximizes $\tilde{u}(p,\tau,\Gamma)$ knowing that $p\in[0,P]\cap[p_{th}(\tau),\infty)$ and $\tau\in(0,1)$. By fixing τ first and optimizing with respect to p, we have that $u(p,\tau,\Gamma)$ is increasing in p and hence, the optimal power equals P and the value of τ will be constrained by $P\ge p_{th}(\tau)$ or equivalently $\tau\le\tau_{th}(P)$. This analysis is identical to the analysis of the NE and one possible SE solution is the NE in Theorem 1.

At the SE, the legitimate user will choose one of the two possibilities which provides a higher SKG utility. From Proposition 2, we know that the NJ state cannot provide a strictly higher utility than the NE state. Hence, whenever $\tau^{NE}<\tau_{th}(P)$, the utility of the unique NE is strictly higher than that of the NJ state. This implies a unique SE that is identical to the NE. If $\tau^{NE}=\tau_{th}(P)$, this means that $(p^{NJ},\tau^{NJ})=(P,\tau^{NE})$ which implies that the utilities at both states NJ and NE are identical. Both the NE (in Theorem 1) and NJ (in Proposition 1) states are SE solutions: $(P,\tau_{th}(P),\Gamma)$ and $(P,\tau_{th}(P),0)$.

The remaining question is whether there exist other solutions when player L chooses the strategy $(p^{SE},\tau^{SE})=(P,\tau_{th}(P))$. In this case, the jammer is rendered indifferent between all of its actions $\gamma\in[0,\Gamma]$, which means that it is also rendered unpredictable. As opposed to the NE, the SE requires the legitimate user to be able to anticipate precisely the jammer's response. To avoid this problem, the leader can silence the jammer by transmitting with power $p=P-\varepsilon$ or ensures that the jammer transmits with full power by transmitting at power $p=P+\varepsilon$, where ε could be made arbitrarily small and, hence, has no practical impact. None of the other NE in Theorem 2 can be SEs, since the jammer's response cannot be predictable.

In conclusion, if $\tau^{NE}<\tau_{th}(P)$, then the SE is unique and identical to the NE in Theorem 2. Otherwise, both the NE and the NJ states are SE solutions.

D. Proof of Proposition 3

Assume by absurdum and WLOG that player J has an NE strategy such that the first channel is left unused $\beta^*=(0,\beta_2,\ldots,\beta_{N+1})$ while other channels are used $\beta_i>0$ for some $2\le i\le N$. Exploiting this knowledge, player L will only employ channel hopping on channel 1 and maybe spreading with non zero probability at the NE. To see this, we write the expected payoff of player L assuming $\beta_1=0$

$$\begin{aligned}&2N\hat{u}(\alpha^*,\beta^*)\\ &=\sum_{i=2}^{N}\{\alpha_i(1-\beta_i-\beta_{N+1})C(NP,0)\\ &\quad+\alpha_i\beta_i C(NP,N\Gamma)+\alpha_i\beta_{N+1}C(NP,\Gamma)\}\\ &\quad+\alpha_{N+1}(1-\beta_{N+1})[(N-1)C(P,0)+C(P,N\Gamma)]\\ &\quad+\alpha_{N+1}\beta_{N+1}NC(P,\Gamma)\alpha_1(1-\beta_{N+1})(N-1)C(NP,0).\end{aligned}$$

Since $C(NP,0)>C(NP,N\Gamma)$ and there exists some $\beta_i>0$, we have that:

$$\begin{aligned}&(1-\beta_{N+1})(N-1)C(NP,0)\\ &\qquad>\sum_{i\neq1}[\beta_i C(NP,N\Gamma)+(1-\beta_i-\beta_{N+1})C(NP,0)].\end{aligned}$$

This means that, if the jammer does not use channel 1, the legitimate ndes will only employ this channel and none of the other channel hopping strategies and the NE will be of the form $\alpha^*=(1-\alpha_{N+1},0,\ldots,0,\alpha_{N+1})$. The utility becomes:

$$\begin{aligned}&2N\hat{u}(\alpha^*,\beta^*)\\ &\quad=(1-\alpha_{N+1})(1-\beta_{N+1})(N-1)C(NP,0)\\ &\qquad+\alpha_{N+1}(1-\beta_{N+1})[(N-1)C(P,0)+C(P,N\Gamma)]\\ &\qquad+\alpha_{N+1}\beta_{N+1}NC(P,\Gamma).\end{aligned}$$

But now, if the jammer uses all channel hopping probabilities back in channel 1, he can strictly decrease the utility. Assume that the jammer switches from the initial β^* to $\delta=(1-\beta_{N+1},0,\ldots,0,\beta_{N+1})$. The payoff becomes:

$$\begin{aligned}&2N\hat{u}(\alpha^*,\delta)\\ &\quad=(1-\alpha_{N+1})(1-\beta_{N+1})(N-1)C(NP,N\Gamma)\\ &\qquad+\alpha_{N+1}(1-\beta_{N+1})[(N-1)C(P,0)+C(P,N\Gamma)]\\ &\qquad+\alpha_{N+1}\beta_{N+1}NC(P,\Gamma).\end{aligned}$$

Since $\hat{u}(\alpha^*, \beta^*) > \hat{u}(\alpha^*, \delta)$, the jammer has an incentive to deviate from the NE which is a contradiction. Thus, the jammer uses either all or none of the channel hopping actions. For player L, the proof follows similarly.

E. Proof of Proposition 4

Let us write the linear equations obtained when the players are indifferent among their channel hopping actions. There are four very similar cases depending on whether the players use spread with zero probability at the NE or not. We only detail one case here below. If both players use spread at the NE, the following conditions must be met:

$$\begin{aligned}&\alpha_i C(NP, N\Gamma) + (1 - \alpha_i - \alpha_{N+1})C(NP, 0)\\ &\qquad + \alpha_{N+1}[(N-1)C(P,0) + C(P, N\Gamma)] = c_\alpha,\\ &\beta_i C(NP, N\Gamma) + (1 - \beta_i - \beta_{N+1})C(NP, 0)\\ &\qquad + \beta_{N+1} C(NP, \Gamma) = c_\beta.\end{aligned}$$

The equations in α illustrate that player J becomes indifferent among their pure channel hopping actions at the NE. Similarly, the equations in β make player L indifferent among their pure channel hopping actions at the NE. We remark that all these equations are identical in the sense that their coefficients do not depend on the index i of the α and β variables. This means that their solutions are of the form: $\alpha_i = a$ and $\beta_i = b$ for any $i \leq N$. Therefore – irrespective of whether the players employ or not spreading at the NE – if both players employ the channel hopping strategy, then the NE takes on the special form $\alpha^* = (a, \ldots, a, (1 - Na))$, $\beta^* = (b, \ldots, b, (1 - Nb))$ for some $0 \leq a \leq 1/N$, $0 \leq b \leq 1/N$.

F. Proof of Theorem 4

If $N = 1$, the NE analysis is trivial and both players transmit at full powers $(NP, N\Gamma)$. If $N > 1$ and given the strict convexity of $C(p, \gamma)$ in γ, we have the following inequality for all p, $\gamma_1 \neq \gamma_2$ and $\lambda \in (0, 1)$:

$$C(p, \lambda\gamma_1 + (1-\lambda)\gamma_2) < \lambda C(p, \gamma_1) + (1-\lambda)C(p, \gamma_2).$$

By taking $p = P$, $\gamma_1 = 0, \gamma_2 = N\Gamma, \lambda = \frac{N-1}{N}$, we obtain:

$$NC(P, \Gamma) < (N-1)C(P, 0) + C(P, N\Gamma) \tag{32}$$

Similarly, by taking $p = NP$, $\gamma_1 = 0, \gamma_2 = N\Gamma, \lambda = \frac{N-1}{N}$, we obtain:

$$NC(NP, \Gamma) < (N-1)C(NP, 0) + C(NP, N\Gamma). \tag{33}$$

From Proposition 3 and Proposition 4, the NE can only take nine forms which are not mutually exclusive. Each case is studied using Definition 1 and for which necessary and sufficient conditions are provided. Then, using (32) and (33), we show that only three of the nine cases can occur. The proof is rather long and tedious and only a sketch containing the main ideas is provided below. 1) *Both players spread at the NE* (i.e., $\alpha^* = \beta^* = e_{N+1}$), iff $C(NP, \Gamma) < NC(P, \Gamma)$ and $(N-1)C(P, 0) + C(P, N\Gamma) > NC(P, \Gamma)$. The second condition is always true due to (32).

2) *Both players use only channel hopping at the NE* (i.e., $\alpha^* = \beta^* = (1/N, \ldots, 1/N, 0)$), iff $C(NP, N\Gamma) + (N-1)C(NP, 0) > N(N-1)C(P, 0) + NC(P, N\Gamma)$ and $C(NP, N\Gamma) + (N-1)C(NP, 0) < NC(NP, \Gamma)$. This case is impossible because of (33).

3) *The game has a strictly mixed NE*, i.e., all actions are used with non-zero probability, of the form $\alpha^* = (a, \ldots, a, (1 - Na))$, $\beta^* = (b, \ldots, b, (1 - Nb))$ iff there exist $0 < a < 1/N$ and $0 < b < 1/N$ such that both players are indifferent among all their pure strategies. Let us write the condition for $(a, \ldots, a, 1 - Na)$ to be a NE and for which the jammer is indifferent among their pure strategies by Definition 1. This yields the following linear equation:

$$\begin{aligned}&a[NC(NP, \Gamma) - C(NP, N\Gamma) - (N-1)C(NP, 0)]\\ &\qquad = (1 - Na)[(N-1)C(P, 0) + C(P, N\Gamma) - NC(P, \Gamma)],\end{aligned}$$

where the term on the LHS is a strictly negative value from $a > 0$ and (33) and the RHS is a strictly positive value from $a < 1/N$ and (32). Thus, this case can never occur.

4) *Player L only channel hops and player J uses both channel hopping and spreading at the NE*: $\alpha^* = (1/N, \ldots, 1/N, 0)$ and $\beta^* = (b, \ldots, b, (1 - Nb))$, iff $C(NP, N\Gamma) + (N-1)C(NP, 0) = NC(NP, \Gamma)$, $0 < b < 1/N$, and $Nb[(N-1)C(P, 0) + C(P, N\Gamma)] + (1 - Nb)NC(P, \Gamma) < bC(NP, N\Gamma) + (N-1)bC(NP, 0) + (1 - Nb)C(NP, \Gamma)$, where b is chosen such that player L is indifferent among their pure strategies. Given (33), the above equality never holds.

5) *Player J only channel hops and player L uses both channel hopping and spreading at the NE* (i.e., $\alpha^* = (a, \ldots, a, (1 - Na))$ and $\beta^* = (1/N, \ldots, 1/N, 0))$, iff $C(NP, N\Gamma) + (N-1)C(NP, 0) = N(N-1)C(P, 0) + C(P, N\Gamma)$, $0 < a < 1/N$, and $MaC(NP, \Gamma) + (1 - Na)NC(P, \Gamma) > aC(NP, N\Gamma) + (N-1)aC(NP, 0) + (1 - Na)[(N-1)C(P, 0) + C(P, N\Gamma)]$ where a is chosen such that player J is indifferent among their pure strategies. The last inequality condition becomes:

$$\begin{aligned}&a[NC(NP, \Gamma) - C(NP, N\Gamma) - (N-1)C(NP, 0)]\\ &\qquad > (1 - Na)[(N-1)C(P, 0) + C(P, N\Gamma) - NC(P, \Gamma)]\end{aligned}$$

where the term on the LHS is a strictly negative value from $a > 0$ and (33) and the RHS is a strictly positive value from $a < 1/N$ and (32). Thus, this case can never occur.

6) *Player L spreads and player J channel hops at the NE* (i.e., $\alpha^* = e_{N+1}$ and $\beta^* = (\beta_1, \ldots, \beta_N, 0)$), iff $NC(P, \Gamma) > (N-1)C(P, 0) + C(P, N\Gamma)$, $NC(NP, 0) - N(N-1)C(P, 0) - NC(P, N\Gamma) < C(NP, 0) - C(NP, N\Gamma)$ and β_i meet some additional constraints. Because of (32) this case never occurs as the first condition is never satisfied.

7) *Player J spreads and player L channel hops at the NE* (i.e., $\beta^* = e_{N+1}$ and $\alpha^* = (\alpha_1, \ldots, \alpha_N, 0)$), iff $C(NP, \Gamma) > NC(P, \Gamma)$ and $NC(NP, 0) - NC(NP, \Gamma) > C(NP, 0) - C(NP, N\Gamma)$. The NE strategies of player L are given by the (infinite number) of solutions to the following system of linear inequalities:

$$\begin{cases} 0 \leq \alpha_i \leq 0, \ \forall i, \ \sum_{j=1}^{N} \alpha_j = 1 \\ \alpha_i < \frac{C(NP,0) - C(NP,\Gamma)}{C(NP,0) - C(NP,N\Gamma)}, \ \forall i \leq N. \end{cases}$$

The second condition is always true (33). From (33), the above system of inequality always has the uniform probability over the channels solution $\alpha^* = (1/N, \ldots, 1/N, 0)$.

8) *Player L spreads and player J employs all their actions at the NE* (i.e., $\alpha^* = e_{N+1}$, $\beta^* = (\beta_1, \ldots, \beta_{N+1})$), iff $(N-1)C(P,0) + C(P, N\Gamma) = NC(P,\Gamma)$ and $\beta_i, \forall i$ meet some additional constraints that are not detailed here. The reason is that, given (32), the equality condition never holds and, hence, this case is impossible.

9) *Player J spreads and player L employs all their actions at the NE* (i.e., $\beta^* = e_{N+1}$ and $\alpha^* = (\alpha_1, \ldots, \alpha_N, \alpha_{N+1})$), iff $C(NP,\Gamma) = NC(P,\Gamma)$ and the solutions to the following linear system of inequalities are NE strategies for player L:

$$\begin{cases} 0 \le \alpha_i \le 1, \ \forall i, \ \sum_{j=1}^{N} \alpha_j = 1 \\ \alpha_i[C(NP, N\Gamma) - C(NP,0)] + \alpha_{N+1}[(N-1)C(P,0) \\ +C(P,N\Gamma) - C(NP,0) + C(NP,\Gamma) - NC(P,\Gamma)] \\ > C(NP,\Gamma) - C(NP,0), \ \forall i \le N. \end{cases}$$

Notice that, by taking $\alpha_{N+1} = 0$, the above system of linear equations is precisely the one in case 7 which has an infinite number of solutions, and in particular $\alpha_i = 1/N, \ \forall i \le N$. Similarly, $\alpha_i = 0$ for all $i \le N$ and $\alpha_{N+1} = 1$ (player L spreads) is also a solution, which follows directly from (32).

G. Proof of Proposition 5

The best-response for the jammer is defined as $\beta^{BR}(\alpha) = \arg\min_\beta \hat{u}(\alpha,\beta)$, where $\beta^{BR}(\alpha)$ represents the best action the jammer can take knowing that the legitimate player choses α. The payoff is affine in β and can be rewritten it as $\hat{u}(\alpha,\beta) = \sum_{i=1}^{N+1} \beta_i c_i(\alpha) + c_0(\alpha)$, with the coefficients:

$$\begin{aligned} c_i(\alpha) &= \alpha_i[C(NP,N\Gamma) - C(NP,0)] \\ &\quad + \alpha_{N+1}[C(P,N\Gamma) - C(P,0)], \quad i \le N, \\ c_{N+1}(\alpha) &= \sum_{j=1}^{N} \alpha_i[C(NP,\Gamma) - C(NP,0)] \\ &\quad + N\alpha_{N+1}[C(P,\Gamma) - C(P,0)], \\ c_0(\alpha) &= \sum_{j=1}^{N} \alpha_i C(NP,0) + N\alpha_{N+1} C(P,0). \end{aligned} \tag{34}$$

Thus, we observe that to find the best-response function $\beta^{BR}(\alpha)$, the jammer has to solve a linear program under the constraints: $\beta_i \ge 0, \ \forall\, i$ and $\sum_{j=1}^{N+1} \beta_j = 1$. The SE action of the leader, anticipating that the jammer will best respond to their own choice is given by:

$$\begin{aligned} \alpha^{SE} &= \arg\max_\alpha \ \hat{u}(\alpha, \beta^{BR}(\alpha)) \\ &= \arg\max_\alpha \ \left\{ \min_{j>0} c_j(\alpha) + c_0(\alpha) \right\}. \end{aligned}$$

Player L can anticipate the response of the jammer, who seeks to minimize the coefficients $c_j(\alpha)$. We remark that: $c_{N+1}(\alpha) = (1-\alpha_{N+1})[C(NP,\Gamma) - C(NP,0)] + N\alpha_{N+1}[C(P,\Gamma) - C(P,0)]$ and $c_0(\alpha) = (1-\alpha_{N+1})C(NP,0) + N\alpha_{N+1}$ do not depend on the way in which the load $1-\alpha_{N+1}$ is spread over the channel hopping actions. Therefore we can only focus on $c_i(\alpha), \ 1 \le i \le N$.

If player L uses channel hopping strategies with uniform probability $\alpha^{(1)} = (a, \ldots, a, 1-Na)$, all coefficients will be equal $c_i(\alpha^{(1)}) = a[C(NP,N\Gamma) - C(NP,0)] + (1-Na)[C(P,N\Gamma) - C(P,0)]$. This means that the jammer is indifferent between the different channels $\min_{1\le j\le N} c_j(\alpha) = a[C(NP,N\Gamma) - C(NP,0)] + (1-Na)[C(P,N\Gamma) - C(P,0)]$.

Now, if player L has a preference for a certain channel, say for channel 1: $\alpha^{(2)} = (a+\delta_1, a-\delta_2, \ldots, a-\delta_N, 1-Na)$, with $\sum_{j=2}^{N} \delta_j = \delta_1 > 0$, the coefficients will be: $c_1(\alpha^{(2)}) = (a+\delta)[C(NP,N\Gamma) - C(NP,0)] + (1-Na)[C(P,N\Gamma) - C(P,0)]$, $c_i(\alpha^{(2)}) = (a-\delta_i)[C(NP,N\Gamma) - C(NP,0)] + (1-Na)[C(P,N\Gamma) - C(P,0)]$. In this case, the jammer will profit from this information and will put all their channel hopping load on channel 1 alone: $\beta_1^{BR}(\alpha^{(2)}) = 1 - \beta_{N+1}^{BR}(\alpha^{(2)})$, $\beta_i(\alpha^{(2)}) = 0, \forall 2 \le N$ and $\min_{1\le j\le N} c_j(\alpha^{(2)}) = (a+\delta)[C(NP,N\Gamma) - C(NP,0)] + (1-Na)[C(P,N\Gamma) - C(P,0)]$. But this means that $\min_{1\le j\le N} c_j(\alpha^{(2)}) < \min_{1\le j\le N} c_j(\alpha^{(1)})$, which further implies that $\hat{u}(\alpha^{(1)}, \beta^{BR}(\alpha^{(1)})) < \hat{u}(\alpha^{(2)}, \beta^{BR}(\alpha^{(2)}))$. This means that player L will lose in utility by not assigning uniform probability to the channel hopping strategies.

H. Proof of Theorem 5

Proposition 1 tells us that the SE strategy of player L is of the form: $\alpha^{SE} = (a, \ldots, a, (1-Na))$ for some $a \in [0, 1/N]$, which is to be determined. The coefficients in (34) become:

$$\begin{aligned} c_i(\alpha^{SE}) &= a[C(NP,N\Gamma) - C(NP,0)] \\ &\quad + (1-Na)[C(P,N\Gamma) - C(P,0)], \quad i \le N \\ c_{N+1}(\alpha^{SE}) &= Na[C(NP,\Gamma) - C(NP,0)] \\ &\quad + N(1-Na)[C(P,\Gamma) - C(P,0)]. \end{aligned}$$

Using the fact that $C(p,\gamma)$ is convex w.r.t. γ for a fixed p, we have the following inequalities: $NC(P,\Gamma) < (N-1)C(P,0) + C(P,N\Gamma)$ and $NC(NP,\Gamma) < (N-1)C(NP,0) + C(NP,N\Gamma)$ which imply that $c_i(\alpha^{SE}) < c_{N+1}(\alpha^{SE})$. This means that the jammer's strategy is to spread always: $\beta^{SE} = (0, \ldots, 0, 1)$. The SE utility becomes:

$$\hat{u}(\alpha^{SE}, \beta^{SE}) = aC(NP,\Gamma) + (1-Na)C(P,\Gamma). \tag{35}$$

This implies that, if $C(NP,\Gamma) > NC(P,\Gamma)$ player L will only channel hop with uniform probability $a = 1/N$. If $C(NP,\Gamma) < NC(P,\Gamma)$ player L will only spread $a = 0$. If $C(NP,\Gamma) = NC(P,\Gamma)$ then the legitimate user is indifferent between spreading and channel hopping and all $a \in [0, 1/N]$ are solutions.

REFERENCES

[1] E. V. Belmega and A. Chorti, "Energy harvesting in secret key generation systems under jamming attacks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017.

[2] A. Chorti and E. V. Belmega, "Secret key generation in Rayleigh block fading AWGN channels under jamming attacks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–7.

[3] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography—Part I: Secret sharing," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1121–1132, Jul. 1993.

[4] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.

[5] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography—Part II: CR capacity," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 225–240, Jan. 1998.
[6] C. Bennett, G. Brassard, C. Crépeau, and U. Maurer, "Generalized privacy amplification," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1915–1923, Nov. 1995.
[7] C. Ye, A. Reznik, and Y. Shah, "Extracting secrecy from jointly Gaussian random variables," in *Proc. Int. Symp. Inf. Theory (ISIT)*, Seattle, WA, USA, Jul. 2006, pp. 2593–2597.
[8] C. Ye, S. Mathur, A. Reznik, Y. Shah, W. Trappe, and N. B. Mandayam, "Information-theoretically secret key generation for fading wireless channels," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 2, pp. 154–240, Jun. 2010.
[9] T.-H. Chou, S. C. Draper, and A. M. Sayeed, "Key generation using external source excitation: Capacity, reliability, and secrecy exponent," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2455–2474, Apr. 2012.
[10] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1550–1573, Aug. 2014.
[11] U. Maurer and S. Wolf, "Secret-key agreement over unauthenticated public channels—Part III: Privacy amplification," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 839–851, Apr. 2003.
[12] V. Yakovlev, V. Korzhik, and G. Morales-Luna, "Key distribution protocols based on noisy channels in presence of an active adversary: Conventional and new versions with parameter optimization," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2535–2549, Jun. 2008.
[13] C. Saiki and A. Chorti, "A novel physical layer authenticated encryption protocol exploiting shared randomness," in *Proc. IEEE Conf. Commun. Netw. Secur. (CNS)*, Sep. 2015, pp. 113–118.
[14] M. Zafer, D. Agrawal, and M. Srivatsa, "Limitations of generating a secret key using wireless fading under active adversary," *IEEE/ACM Trans. Netw.*, vol. 20, no. 5, pp. 1440–1451, Oct. 2012.
[15] R. Molière, F. Delaveau, C. L. K. Ngassa, C. Lemenager, T. Mazloum, and A. Sibille, "Tag signals for early authentication and secret key generation in wireless public networks," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Jun. 2015, pp. 108–112.
[16] G. T. Amariucai, S. Wei, and R. Kannan, "Gaussian jamming in block-fading channels under long term power constraints," in *Proc. Int. Symp. Inf. Theory (ISIT)*. Nice, France Jun. 2007, pp. 1001–1005.
[17] X. Song, P. Willett, S. Zhou, and P. Luh, "The MIMO radar and jammer games," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 687–699, Feb. 2012.
[18] S. Wei, R. Kannan, V. Chakravarthy, and M. Rangaswamy, "CSI usage over parallel fading channels under jamming attacks: A game theory study," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 1167–1175, Apr. 2012.
[19] R. El-Bardan, S. Brahma, and P. Varshney, "Strategic power allocation with incomplete information in the presence of a jammer," *IEEE Trans. Commun.*, vol. 64, no. 8, pp. 3467–3479, Aug. 2016.
[20] L. Xiao, T. Chen, J. Liu, and H. Dai, "Anti-jamming transmission Stackelberg game with observation errors," *IEEE Commun. Lett.*, vol. 19, no. 6, pp. 949–952, Jun. 2015.
[21] J. Guo, N. Zhao, F. R. Yu, X. Liu, and V. C. M. Leung, "Exploiting adversarial jamming signals for energy harvesting in interference networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 1267–1280, Feb. 2017.
[22] H. Fang, L. Xu, and K.-K. R. Choo, "Stackelberg game based relay selection for physical layer security and energy efficiency enhancement in cognitive radio networks," *Appl. Math. Comput.*, vol. 296, pp. 153–167, Mar. 2017.
[23] A. Mukherjee and A. L. Swindlehurst, "Jamming games in the MIMO wiretap channel with an active eavesdropper," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 82–91, Jan. 2013.
[24] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge, U.K: Cambridge Univ. Press, 2011.
[25] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications Handbook. Electronic*. New York, NY, USA: McGraw-Hill, 2002.
[26] R. Poisel, *Modern Communications Jamming Priniples and Techniques*. Norwood, MA, USA: Artech House, 2003.
[27] M. Strasser, C. Pöpper, S. Čapkun, and M. Čagalj, "Jamming-resistant key establishment using uncoordinated frequency hopping," in *Proc. IEEE Symp. Secur. Privacy*, May 2008, pp. 64–78.
[28] C. Pöpper, M. Strasser, and S. Čapkun, "Anti-jamming broadcast communication using uncoordinated spread spectrum techniques," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 5, pp. 703–715, Jun. 2010.
[29] R. Rajesh, V. Sharma, and P. Viswanath, "Capacity of Gaussian channels with energy harvesting and processing cost," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2563–2575, May 2014.
[30] Y. Gu and S. Aïssa, "RF-based energy harvesting in decode-and-forward relaying systems: Ergodic and outage capacities," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6425–6434, Nov. 2015.
[31] S. Ulukus *et al.*, "Energy harvesting wireless communications: A review of recent advances," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 360–381, Mar. 2015.
[32] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA, USA: MIT Press, 1991.
[33] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[34] J. V. Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton, NJ, USA: Princeton Univ. Press, 2007.
[35] J. Xu, L. Duan, and R. Zhang, "Proactive eavesdropping via jamming for rate maximization over Rayleigh fading channels," *IEEE Wireless Commun. Lett.*, vol. 5, no. 1, pp. 80–83, Feb. 2016.

E. Veronica Belmega (S'08–M'10) received the M.Sc. (engineer diploma) degree from the University Politehnica of Bucharest, Romania, in 2007, the M.Sc. and Ph.D. degrees from the Université Paris-Sud 11, France, in 2007 and 2010, respectively. From 2010 to 2011, she was a Post-Doctoral Researcher in a joint project between Princeton University, USA, and the Alcatel-Lucent Chair on Flexible Radio in Supélec, France. She is currently an Assistant Professor with ETIS, UMR 8051, Université Paris Seine, Université Cergy-Pontoise, ENSEA, CNRS, France, and Inria. She was one of the ten recipients of the L'Oréal-UNESCO-French Academy of Science Fellowship: For young women doctoral candidates in science in 2009.

Arsenia Chorti (S'00–M'05) received the M.Eng. degree in electrical and electronics engineering from the University of Patras, Greece, the D.E.A. degree in electronics from the University Pierre et Marie Curie-Paris VI, France, and the Ph.D. degree in electrical engineering from Imperial College London, U.K. She holds the post-doctoral positions with the Universities of Southampton, U.K., TCU, Greece, and UCL, U.K., from 2005 to 2008. She has served as a Senior Lecturer in communications with Middlesex University, U.K., from 2008 to 2010. From 2010 to 2013, she was a Marie Curie IOF Researcher with Princeton University, USA, and ICS-FORTH, Greece. Since 2013, she has been holding a lecturer position in communications and networks with CSEE University Essex, U.K. She is currently a Visiting Research Collaborator with Princeton University.

Learning to Be Green: Robust Energy Efficiency Maximization in Dynamic MIMO–OFDM Systems

Panayotis Mertikopoulos, *Member, IEEE*, and E. Veronica Belmega, *Member, IEEE*

***Abstract*—In this paper, we examine the maximization of energy efficiency (EE) in next-generation multiuser MIMO–OFDM networks that vary dynamically over time—e.g., due to user mobility, fluctuations in the wireless medium, modulations in the users' load, etc. Contrary to the static/stationary regime, the system may evolve in an arbitrary manner, so users must adjust "on the fly," without being able to predict the state of the system in advance. To tackle these issues, we propose a simple and distributed online optimization policy that leads to *no regret*, i.e., it allows users to match (and typically outperform) even the best fixed transmit policy in hindsight, irrespective of how the system varies with time. Moreover, to account for the scarcity of perfect channel state information (CSI) in massive MIMO systems, we also study the algorithm's robustness in the presence of measurement errors and observation noise. Importantly, the proposed policy retains its no-regret properties under very mild assumptions on the error statistics: on average, it enjoys the same performance guarantees as in the noiseless deterministic case. Our analysis is supplemented by extensive numerical simulations, which show that, in realistic network environments, users track their individually optimum transmit profile even under rapidly changing channel conditions, achieving gains of up to 600% in energy efficiency over uniform power allocation policies.**

***Index Terms*—Energy efficiency, imperfect CSI, MIMO, OFDM, no regret, online optimization.**

I. Introduction

THE WILDFIRE spread of Internet-enabled mobile devices and the exponential growth of bandwidth-hungry applications is putting existing wireless systems under enormous strain and is one of the driving forces behind the transition to fifth generation (5G) mobile networks [1]. The ICT industry is thus faced with a formidable mission: data rates must be increased significantly so as to meet the soaring demand for wireless broadband, but this task must be accomplished under an extremely tight energy budget. With this in mind, current design requirements for 5G systems target a dramatic decrease in energy-per-bit consumption of the order of $1,000\times$ or more [2], [3].

Manuscript received April 15, 2015; revised September 7, 2015; accepted December 11, 2015. Date of publication March 21, 2016; date of current version May 11, 2016. This work was supported in part by the French National Research Agency project NETLEARN (ANR–13–INFR–004) and in part by the ENSEA, Cergy-Pontoise, France. This work was presented in part at VTC2015-Spring.

P. Mertikopoulos is with the French National Center for Scientific Research (CNRS), Laboratoire d'Informatique de Grenoble (LIG), Grenoble F-38000, France and also with Inria, France (e-mail: panayotis.mertikopoulos@imag.fr).

E. V. Belmega is with ETIS/ENSEA-UCP-CNRS, Cergy-Pontoise, France, and also with Inria, France.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSAC.2016.2544600

A contending technology to achieve these design targets is the emerging massive MIMO (multiple-input and multiple-output) paradigm. Coupled with the use of multiple carrier frequencies via orthogonal frequency division multiplexing (OFDM), massive MIMO "goes large" by employing inexpensive service antennas to focus energy into ever smaller regions of space [4]–[6]. As a result, very large MIMO arrays can greatly enhance the reliability of wireless connections and increase throughput and energy efficiency (EE) by a factor of $10\times$ to $100\times$ without requiring the deployment of expensive new air interfaces [1], [6]. However, due to the massive complexity and variability of such systems, a crucial challenge that arises is that wireless users must also be capable of adapting to a dynamic spectrum landscape "on the fly", typically with minimal coordination and limited information at the device end.

An added challenge in the above considerations is that wireless users often do not have access to perfect CSI and co-channel interference (CCI) measurements, especially at the transmitter end – for instance, due to pilot contamination in massive MIMO systems [6]. In particular, if the system operates in the presence of uncertainty (imperfect CSI, observation noise, etc.), optimization techniques that rely on a greedy, "one-off" calculation of optimal transmit characteristics (such as water-filling) are no longer suitable because stochastic fluctuations could lead the system to a suboptimal state [7], [8]. On that account, our main objective in this paper will be to provide an adaptive transmit policy for energy efficiency maximization in dynamic MIMO–OFDM networks that are subject to uncertainty, feedback errors and/or other unpredictable changes in the wireless medium.

In the general context of MIMO–OFDM systems, the vast majority of works on energy efficiency maximization and energy-efficient power allocation have focused on two limit cases [9]. In the static regime [10]–[15], the attributes of the wireless system under study (channel gains, user load, etc.) are assumed effectively static and the system's analysis revolves around techniques from the theory of non-cooperative games and optimization (continuous or discrete). At the other end of the spectrum, in the ergodic regime [12], [16], the wireless medium is assumed to evolve over a very fast time scale, typically following a sequence of independent and identically distributed (i.i.d.) random variables; consequently, the figure of merit in problems of this type is the ratio between the stochastic average of the users' rate and their power consumption. All these works study the trade-off between the Shannon achievable rate and power consumption either for a single user (via fractional programming) or multiple ones (using the theory of non-cooperative games). Finally, in the static channel regime,

[17]–[22] consider a throughput model that depends on the connection's bit error rate (BER) and use tools from game theory to characterize the system's equilibrium states.

In this paper, we focus squarely on dynamic MIMO–OFDM systems that evolve *arbitrarily* over time (e.g. due to channel variability, fading, mobility, etc.), and we make no statistical hypotheses regarding the dynamics that govern the network's evolution (such as stationarity or ergodicity). As opposed to the stationary/ergodic regime discussed above, static solution concepts are no longer relevant because there is no underlying target state to attain (either static or in the mean); as such, no conclusions can be drawn from the existing literature on energy-efficient power allocation. Instead, users have to optimize their transmit characteristics on the fly, based only on locally available information of the past state of the system, and hoping to track (or at least emulate the performance of) the *a posteriori* optimum transmit policy.

The most widely used optimization criterion in this setting is that of *regret minimization*, a seminal notion which was first introduced by Hannan [23] and which has since given rise to a vigorous literature at the interface of machine learning, optimization, statistics, and game theory – for a comprehensive survey, see e.g. [24]–[26]. More precisely, in the language of game theory, a user's regret over a given time horizon is simply the difference between his average payoff (over the time horizon in question) and the payoff that he would have obtained if he had employed the best possible fixed action in hindsight. Accordingly, in our case, regret minimization corresponds to dynamic transmit policies that are asymptotically optimal in hindsight, irrespective of how the users' effective wireless medium evolves over time.

A regret-based approach was recently employed by the authors of [27] who studied the problem of power control in infrastructureless wireless networks and proposed an algorithm that minimizes the users' (internal) regret to attain the system's equilibrium. In a similar vein, [28] studied the transient phase of the Foschini–Miljanic (FM) power control algorithm in static environments and used the notion of *swap regret* [29] to propose alternative convergent power control schemes; even more recently, [30] showed that the FM dynamics lead to no regret, so they retain their optimality properties in dynamic environments. Finally, [31] and [32] used online optimization techniques and a methodology based on matrix exponential learning [7], [8], [33], [34] to derive a no-regret adaptive transmit policy for power control and throughput maximization in cognitive radio networks respectively. However, the proposed policies drive wireless users to transmit at either full or minimum power (subject to their rate requirements), so they cannot be applied to minimize energy-per-bit consumption in dynamic MIMO–OFDM systems.

Summary of results and paper outline: In this paper, we formulate the maximization of energy efficiency in dynamic MIMO–OFDM systems as an online optimization problem, and we draw on Zinkevich's seminal online gradient ascent (OGA) methodology [35] to derive an adaptive transmit policy that leads to no regret. In particular, we show that the proposed algorithm guarantees an $\mathcal{O}(T^{-1/2})$ regret bound after T update epochs (transmission frames), and this bound tightens to $\mathcal{O}(\log T/T)$ if the users' channel gains always remain above a given level. Furthermore, to address the lack of perfect measurements and channel state information at the transmitter (CSIT), we show that the proposed algorithm retains these properties under very mild statistical hypotheses that are satisfied by the vast majority of error distributions. Specifically, as long as *a)* there is no *systematic* error in the measurement process; and *b)* the probability of a very high error z is not higher than $\mathcal{O}(1/z^2)$, the proposed policy leads to no regret and enjoys a mean bound of the same order as in the deterministic setting.

The performance of the proposed transmit policy is validated by means of extensive numerical simulations modeling a cellular orthogonal frequency-division multiple access (OFDMA) network with multiple base stations and mobile MIMO users under realistic wireless propagation, fading and mobility features. Our results show that the proposed policy represents a scalable and flexible method that allows users to attain very high energy efficiency levels, exhibiting gains of up to 600% over uniform/fixed power allocation policies and with surprisingly modest feedback requirements.

This paper greatly extends our recent conference paper [36] where we derived a continuous-time exponential learning method for energy efficiency maximization in dynamic single-input and single-output (SISO) systems. Compared to [36], the current paper provides a bona fide learning algorithm for multiple-antenna systems, with discrete-time updates and performance guarantees, and with proven robustness in the presence of uncertainty and observation noise.

The rest of our paper is structured as follows: in Section II, we present our wireless system model and we formulate the problem of dynamic energy efficiency maximization as an online semidefinite program. In Section III, we derive our online learning policy, and we establish its no-regret properties and performance guarantees under both perfect and imperfect CSI. Finally, our theoretical analysis is supplemented by extensive numerical simulations in Section IV.

II. System Model and Problem Formulation

Consider a wireless network consisting of several point-to-point connections $u \in \mathcal{U} = \{1, \dots, U\}$ (the system's *users*) that are established over a set of orthogonal subcarriers $k \in \mathcal{K} \equiv \{1, \dots, K\}$. Each connection $u \in \mathcal{U}$ comprises a pair of communicating wireless multi-antenna devices with M_u antennas at the transmitter and N_u antennas at the receiver. Thus, if $\mathbf{x}_k^u \in \mathbb{C}^{M_u}$ and $\mathbf{y}_k^u \in \mathbb{C}^{N_u}$ denote the signals transmitted and received over connection u on subcarrier k, we obtain the familiar baseband signal model:

$$\mathbf{y}_k^u = \mathbf{H}_k^{uu}\mathbf{x}_k^u + \sum\nolimits_{u' \neq u} \mathbf{H}_k^{u'u}\mathbf{x}_k^{u'} + \mathbf{z}_k^u, \tag{1}$$

where $\mathbf{H}_k^{u'u} \in \mathbb{C}^{N_u \times M_{u'}}$ denotes the transfer matrix between the u'-th transmitter and the u-th receiver over subcarrier k, while $\mathbf{z}_k^u$ is the ambient noise over the channel (including thermal and atmospheric effects, and modeled as a circularly symmetric Gaussian complex vector). With this in mind, the multi-user interference-plus-noise (MUI) at the intended receiver of the

u-th connection will be:

$$\mathbf{w}_k^u = \sum\nolimits_{u' \neq u} \mathbf{H}_k^{u'u} \mathbf{x}_k^u + \mathbf{z}_k^u, \tag{2}$$

so (1) may be written more simply as:

$$\mathbf{y}_k^u = \mathbf{H}_k^{uu} \mathbf{x}_k^u + \mathbf{w}_k^u. \tag{3}$$

In the rest of this paper, we will focus on a specific connection $u \in \mathcal{U}$ and we will treat the MUI vector $\mathbf{w}_k$ as an aggregate noise variable whose covariance depends on the wireless medium and the transmit characteristics of all other users. Accordingly, if we drop the user index u for notational convenience, the signal model (3) attains the more compact form:

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{w}_k. \tag{4}$$

Hence, assuming Gaussian input and single user decoding (SUD) at the receiver, the Shannon rate at the focal connection will be given by the well-known expression [37]:[1]

$$R(\mathbf{Q}) = \sum\nolimits_{k \in \mathcal{K}} \left[\log\det\left(\mathbf{W}_k + \mathbf{H}_k \mathbf{Q}_k \mathbf{H}_k^\dagger\right) - \log\det \mathbf{W}_k \right], \tag{5}$$

where:

1) $\mathbf{Q}_k = \mathbb{E}[\mathbf{x}_k \mathbf{x}_k^\dagger] \in \mathbb{C}^{M \times M}$ is the user's input signal covariance matrix over subcarrier k.[2]
2) $\mathbf{Q} = \operatorname{diag}(\mathbf{Q}_1, \dots, \mathbf{Q}_K)$ is the power profile of the focal user over all subcarriers.
3) $\mathbf{W}_k = \mathbb{E}[\mathbf{w}_k \mathbf{w}_k^\dagger] \in \mathbb{C}^{N \times N}$ is the MUI covariance matrix of the focal connection (obviously, $\mathbf{W}_k$ depends on all other users in the network).

Remark: The Gaussian input and noise assumptions are fairly standard in the literature: in particular, Gaussian noise is known to be the worst additive noise distribution with respect to the Shannon achievable rate [38] while Gaussian input is optimal against a Gaussian environment [37]. Finally, regarding the decoding technique, SUD has the advantage of being simple, distributed and scalable because it does not require any coordination or signaling among the interfering users.

In view of all this, if we let

$$\widetilde{\mathbf{H}}_k = \mathbf{W}_k^{-1/2} \mathbf{H}_k \tag{6}$$

denote the *effective channel matrix* of the focal user over subcarrier k, the user's Shannon rate (7) can be written more concisely as:

$$R(\mathbf{Q}) = \sum\nolimits_{k \in \mathcal{K}} \log\det\left(\mathbf{I} + \widetilde{\mathbf{H}}_k \mathbf{Q}_k \widetilde{\mathbf{H}}_k^\dagger\right) = \log\det\left(\mathbf{I} + \widetilde{\mathbf{H}} \mathbf{Q} \widetilde{\mathbf{H}}^\dagger\right), \tag{7}$$

where $\widetilde{\mathbf{H}} = \operatorname{diag}(\widetilde{\mathbf{H}}_1, \dots, \widetilde{\mathbf{H}}_K)$ is the block-diagonal sum of the user's effective channel matrices over all subcarriers. Thus, following [11]–[14], the user's *energy efficiency* will be given by:

$$\mathrm{EE}(\mathbf{Q}) = \frac{R(\mathbf{Q})}{P_c + \operatorname{tr}(\mathbf{Q})} = \frac{\log\det\left(\mathbf{I} + \widetilde{\mathbf{H}} \mathbf{Q} \widetilde{\mathbf{H}}^\dagger\right)}{P_c + \operatorname{tr}(\mathbf{Q})}, \tag{8}$$

where $\operatorname{tr}(\mathbf{Q}) = \sum_k \operatorname{tr}(\mathbf{Q}_k)$ is the user's total transmit power while P_c denotes the total power dissipated in all other circuit components of the transmitting device (mixer, frequency synthesizer, digital-to-analog converter, etc.). This efficiency function (which, formally, has units of bits/Joule) has been widely studied in the literature [12], [19], [39] and it captures the fundamental trade-off between higher spectral efficiency and increased battery life. Consequently, in the context of power-limited, energy-aware users, we obtain the maximization problem:

$$\begin{aligned} &\text{maximize} && \mathrm{EE}(\mathbf{Q}), \\ &\text{subject to} && \mathbf{Q} \in \mathcal{Q}, \end{aligned} \tag{9}$$

where

$$\mathcal{Q} = \left\{ \operatorname{diag}(\mathbf{Q}_1, \dots, \mathbf{Q}_K) : \mathbf{Q}_k \succcurlyeq 0,\ \textstyle\sum_k \operatorname{tr}(\mathbf{Q}_k) \leq P_{\max} \right\}, \tag{10}$$

and $P_{\max}$ denotes the user's maximum transmit power.

Of course, the user's energy efficiency function depends not only on the transmitter's signal covariance profile $\mathbf{Q}$, but also on the transmit characteristics of all other users (via the effective channel matrices $\widetilde{\mathbf{H}}_k$). In particular, $\widetilde{\mathbf{H}}$ collects all sources of noise and interference that cannot be controlled by the focal transmit/receive pair, so the user's energy efficiency objective may vary over time in an unpredictable way. On that account, since we wish to focus on dynamic networks that evolve in an arbitrary fashion, we will not be making any specific postulates regarding the behavior of other users in the network or the evolution of the user's actual channel matrix $\mathbf{H}$.

The only generic assumptions that we will make regarding the effective channel matrices $\widetilde{\mathbf{H}}$ are:

(A1) $\widetilde{\mathbf{H}}$ remains bounded in norm over the entire transmission horizon.[3]
(A2) The variability of $\widetilde{\mathbf{H}}$ within each transmission frame is sufficiently slow so that the Shannon mutual information (7) remains a relevant measure of the achievable transmission rate.

With all this at hand, if $\widetilde{\mathbf{H}}(t)$ is the user's effective channel matrix at time t, we obtain the *online* energy efficiency problem:

$$\begin{aligned} &\text{maximize} && \mathrm{EE}(\mathbf{Q}; t), \\ &\text{subject to} && \mathbf{Q} \in \mathcal{Q}, \end{aligned} \tag{OEE}$$

where, in obvious notation:

$$\mathrm{EE}(\mathbf{Q}; t) = \frac{\log\det\left(\mathbf{I} + \widetilde{\mathbf{H}}(t) \mathbf{Q} \widetilde{\mathbf{H}}^\dagger(t)\right)}{P_c + \operatorname{tr}(\mathbf{Q})} \tag{11}$$

denotes the user's energy efficiency function at time t. Thus, given that the user cannot predict the state of the system ahead of time, we will focus on the following sequence of events:

1) At each update period $n = 1, 2, \dots$, the user selects a transmit power profile $\mathbf{Q}(n) \in \mathcal{Q}$.

[1] For the sake of simplicity, constant multiplicative factors such as the bandwidth of the connection have been dropped in (5); these factors are reinstated in the numerical analysis of Section IV.

[2] In the above, expectations are taken over the users' codebooks (assumed Gaussian).

[3] In the standard channel and wireless propagation models used in the literature (Okumura, Hata [40], COST-Hata [41], etc.), the pathloss is caused by the non-zero distance between transmitter and receiver; factors such as RF circuit losses, interference, and shadowing further diminish the user's effective channel gains. The boundedness of $\|\widetilde{\mathbf{H}}\|$ simply reflects these losses.

2) The user's energy efficiency over the current period is determined by the effective channel matrix $\widetilde{\mathbf{H}}(n)$ at the time of transmission.
3) At the end of the period, the user selects a new signal covariance profile $\mathbf{Q}(n+1)$ seeking to maximize his *a priori unknown* objective function $\mathrm{EE}(\mathbf{Q}; n+1)$, and the process repeats.

Of course, the key challenge in this dynamic framework is that the user does not know ahead of time the effective channel matrix $\widetilde{\mathbf{H}}(n+1)$ that determines his energy efficiency function at stage $n+1$, so he must try to adapt to the changing network conditions "on the fly". To be sure, if the user had perfect foresight and knowledge of the evolution of $\widetilde{\mathbf{H}}(n)$ in advance, the (fixed) power profile that maximizes the user's average energy efficiency over a given transmission horizon T would be the solution to the (offline) maximization problem:

$$\max_{\mathbf{Q}\in\mathcal{Q}} \frac{1}{T}\sum\nolimits_{n=1}^{T} \mathrm{EE}(\mathbf{Q}; n). \tag{12}$$

Obviously however, this "oracle" solution cannot be computed without precognitive abilities, so we will focus on adaptive transmit policies $\mathbf{Q}(n)$ that approach the maximal value of (12) asymptotically, *irrespective of the system's evolution over time*.

To make this analysis precise, we follow [24], [25] and we define the user's (cumulative) *regret* at time T as the cumulative difference between the user's achieved EE and the solution of the maximization problem (12), i.e. we let:

$$\mathrm{Reg}(T) = \max_{\mathbf{Q}\in\mathcal{Q}} \sum\nolimits_{n=1}^{T} \left[\mathrm{EE}(\mathbf{Q}; n) - \mathrm{EE}(\mathbf{Q}(n); n)\right]. \tag{13}$$

We then say that a dynamic transmit policy $\mathbf{Q}(n)$ leads to *no regret* if

$$\mathrm{Reg}(T) = o(T), \tag{14}$$

independently of the evolution of the user's energy efficiency function. In this way, a no-regret policy $\mathbf{Q}(n)$ is *asymptotically optimal in hindsight* in that it provides an asymptotic solution to the average energy efficiency maximization problem (12), without requiring any oracle-like capabilities from the user.

The definition of no-regret policies is crucial for the rest of our paper so some remarks are in order:

Remark 1: The seminal notion of regret was first introduced by Hannan [23] and it has since given rise to a vast corpus of research at the interface of optimization, statistics and machine learning—for a recent survey, see e.g. [24], [25]. In particular, if the user's energy efficiency function does not vary with time (i.e. if the user's effective channels are static), standard arguments from the theory of online optimization [24] can be used to show that no-regret policies converge to the solution set $\arg\max_{\mathbf{Q}}\mathrm{EE}(\mathbf{Q})$ of the energy efficiency maximization problem (9).

Of course, if the user had perfect foresight and could predict the future without error, his optimum transmit policy would be given by the *dynamic oracle* solution $\mathbf{Q}^*(n) = \arg\max_{\mathbf{Q}}\mathrm{EE}(\mathbf{Q}; n)$, i.e. the policy that maximizes his energy efficiency at each instance in time (and not only on average). Unfortunately, as is well-known in the theory of online learning, it is not possible to consistently match this dynamic oracle, so $\mathbf{Q}^*(n)$ does not provide an achievable theoretical target [24], [26]. To better understand why, consider a simple example where the user's channel alternates between two values, $\mathbf{H}_a$ and $\mathbf{H}_b$, with corresponding optimal power profiles $\mathbf{Q}_a^*$ and $\mathbf{Q}_b^*$. In this context, the user first selects $\mathbf{Q}(n)$ at stage n and "Nature" subsequently determines the value of the channel ($\mathbf{H}_a$ or $\mathbf{H}_b$), again at stage n. Hence, at each iteration, an adversarial Nature (such as a jammer) could cause the user to underperform with respect to $\mathbf{Q}^*(n)$ by at least $m = \min_{\mathbf{Q}} \max\{|\mathrm{EE}(\mathbf{Q}_a^*; \mathbf{H}_a) - \mathrm{EE}(\mathbf{Q}; \mathbf{H}_a)|, |\mathrm{EE}(\mathbf{Q}_b^*; \mathbf{H}_b) - \mathrm{EE}(\mathbf{Q}; \mathbf{H}_b)|\} > 0$. As a result, *any* adaptive learning policy that relies only on causal information (i.e. no look-ahead into the future) leads to positive regret against the dynamic oracle solution $\mathbf{Q}^*(n)$.

This question can be analyzed by employing more sophisticated regret notions such as that of adaptive [42] or shifting [43] regret. However, since this would require the introduction of significant technical machinery that lies beyond the scope of the current work, we delegate this analysis to future work. For a numerical comparison with the dynamic oracle solution $\mathbf{Q}^*(n)$ in realistic network conditions, we refer the reader to Sec. IV.

Remark 2: We should also note here that the no-regret property (14) is a "worst-case" guarantee that carries no assumptions on the evolution of the user's environment over time: the user's channels could evolve adversarially (e.g. if the user is subject to jamming), or not at all (in the static regime). As such, when the wireless medium is affected only by the behavior of other users in the network, a natural question that arises is whether the use of a no-regret policy by all users leads to an equilibrium of the underlying game.[4] We address this issue in more detail in Section IV.

Remark 3: Another important special case of the online energy efficiency maximization problem (11) is the fast-fading framework where the only variation in the system is due to the channels following a stationary ergodic process – the so-called *ergodic regime* [45], [46]. In this case, assuming full causal knowledge of the channel statistics (but not the channels' actual realization at the time of transmission),[5] the maximization of the user's ergodic EE ratio leads to a stochastic variant of the problem (11) where the user's energy efficiency function is averaged over the channel statistics. By standard arguments from the theory of online optimization [24], it can be shown that a policy that leads to no-regret also solves this averaged problem, so regret minimization is equivalent to ergodic energy efficiency maximization with causal knowledge of channel statistics at the transmitter.

[4]For instance, it is well-known that internal regret minimization implies convergence to the set of correlated equilibria [44].

[5]In the recent paper [12], the authors consider the case where the transmitter has access to instantaneous channel realization information at the time of transmission (zero-delay feedback in continuous time). In this case, ergodic EE optimization is not carried over fixed input covariance matrices but over functional distributions of the form $\tilde{\mathbf{Q}} \equiv \tilde{\mathbf{Q}}(\mathbf{H})$; more precisely, the authors of [12] seek the functional distribution $\tilde{\mathbf{Q}}(\mathbf{H})$ that maximizes the ratio $\mathbb{E}_{\mathbf{H}}[R(\tilde{\mathbf{Q}}(\mathbf{H}))]/\mathbb{E}_{\mathbf{H}}[P_c + \mathrm{tr}(\tilde{\mathbf{Q}}(\mathbf{H}))]$. This leads to a maximization problem defined over an infinite-dimensional Banach space of functional distributions and regret minimization is not equivalent to ergodic energy efficiency maximization in this case; however, since this formulation assumes perfect zero-delay (non-causal) feedback, we do not address it here.

III. Online Learning

A first idea to achieve no regret in the online energy efficiency maximization problem (OEE) would be to calculate at each stage the power profile that maximizes energy efficiency based on the latest available information at the previous stage. However, as can be seen by a standard argument, this policy may lead to *positive* regret: for instance, when the user's channel alternates every other period between two values (say, $\mathbf{H}_a$ and $\mathbf{H}_b$, with corresponding optimal power profiles $\mathbf{Q}_a^*$ and $\mathbf{Q}_b^*$), best-responding to the last observed state performs strictly worse than the *fixed* policy $(\mathbf{Q}_a^* + \mathbf{Q}_b^*)/2$ [24]. With this in mind, we propose in this section an *adaptive* power allocation policy based on Zinkevich's seminal OGA method [35] that utilizes *all* past information in a recursive way.

For simplicity, we first consider the case where the transmitter has access to perfect CSI and MUI measurements, and we derive an anytime $\mathcal{O}(T^{1/2})$ bound for the user's regret. We then show that this bound can be tightened to $\mathcal{O}(\log T)$ if the user's effective channels always remain above a certain threshold and the algorithm's step-size is chosen accordingly. The robustness of these guarantees in the presence of noise and uncertainty is then discussed in Section III-C.

A. Energy Efficiency Maximization as a Concave Problem

The first difficulty in designing a no-regret policy for the online fractional program (OEE) is that the user's energy efficiency function is not concave. This is perhaps most easily seen in the SISO case, in which the user's energy efficiency objective is:

$$\mathrm{EE}(\mathbf{p}) = \frac{\sum_k \log(1 + \tilde{g}_k p_k)}{P_c + \sum_k p_k}, \tag{15}$$

with $\mathbf{p} = (p_1, \dots, p_K)$ denoting the user's power allocation vector and $\tilde{g}_k$ being the effective channel gain of channel k. Clearly, the fractional objective (15) is not concave with respect to any p_k; however, $\mathrm{EE}(\mathbf{p})$ can be recast as a concave function by employing the so-called Charnes–Cooper transformation [47] for turning fractional programs into concave ones.[6] Specifically, if we set

$$x_0 = (P_c + \textstyle\sum_k p_k)^{-1}, \quad \mathbf{x} = x_0 \cdot \mathbf{p}, \tag{16}$$

we readily obtain $\mathrm{EE}(\mathbf{p}) = x_0 \sum_k \log(1 + \tilde{g}_k x_k/x_0)$, and this last function is concave because the summands $x_0 \log(1 + \tilde{g}_k x_k/x_0)$ are jointly concave in x_0 and x_k. We may then drop the parameter x_0 by noticing that $x_0 = \frac{1}{P_c}(1 - \sum_k x_k)$; by rewriting $\mathbf{x}$ as $\mathbf{x} = \mathbf{p}/(P_c + \sum_k p_k)$, solving for $\mathbf{p}$ and substituting in $\mathrm{EE}(\mathbf{p})$, we obtain a concave reformulation of (15).

In a more general MIMO framework, this procedure amounts to the change of variables

$$\mathbf{X} = \frac{P_c + P_{\max}}{P_{\max}} \frac{\mathbf{Q}}{P_c + \mathrm{tr}(\mathbf{Q})}, \tag{17}$$

where we have introduced the normalization constant $(P_c + P_{\max})/P_{\max}$ in order to have $\mathrm{tr}(\mathbf{X}) \leq 1$ for all $\mathbf{Q} \in \mathfrak{Q}$ (with equality if and only if $\mathrm{tr}(\mathbf{Q}) = P_{\max}$). Solving for $\mathbf{Q}$ yields

$$\mathbf{Q} = \frac{P_c P_{\max}}{P_c + P_{\max}(1 - \mathrm{tr}(\mathbf{X}))} \mathbf{X}, \tag{18}$$

so, after substituting in (8), we obtain the objective

$$u(\mathbf{X}) = \frac{P_c + P_{\max}(1 - \mathrm{tr}(\mathbf{X}))}{P_c(P_c + P_{\max})} \log\det\left(\mathbf{I} + \frac{P_c P_{\max} \widetilde{\mathbf{H}} \mathbf{X} \widetilde{\mathbf{H}}^\dagger}{P_c + P_{\max}(1 - \mathrm{tr}(\mathbf{X}))}\right), \tag{19}$$

while the corresponding feasible region of (9) attains the form

$$\mathcal{X} = \left\{\mathrm{diag}(\mathbf{X}_1, \dots, \mathbf{X}_k) : \mathbf{X}_k \succcurlyeq 0 \text{ and } \textstyle\sum_k \mathrm{tr}(\mathbf{X}_k) \leq 1\right\}. \tag{20}$$

Given that $R(\mathbf{Q})$ is concave in $\mathbf{Q}$, the function $F(\mathbf{X}, x) = \frac{P_{\max}}{P_c + P_{\max}} x \cdot R(\mathbf{X}/x)$ is jointly concave in $\mathbf{X}$ and x [48], so $u(\mathbf{X})$ will also be concave in $\mathbf{X}$ as the restriction of $F(\mathbf{X}, x)$ to the convex set $P_c P_{\max} x = P_c + P_{\max}(1 - \mathrm{tr}(\mathbf{X}))$. In this way, (OEE) boils down to the online concave maximization problem:

$$\begin{aligned} &\text{maximize} && u(\mathbf{X}; n), \\ &\text{subject to} && \mathbf{X} \in \mathcal{X}, \end{aligned} \tag{21}$$

where, as before, the dependence on $n = 1, 2 \dots$, reflects the evolution of the user's effective channel matrices over time. In view of all this, we will first derive a no-regret transmit policy $\mathbf{X}(n)$ for the online concave problem (21), and we will then use the inverse transformation (18) to obtain a no-regret policy for (OEE).

B. Learning With Accurate Causal CSI

Building on Zinkevich's online gradient ascent method [35], the core idea of our approach will be to track the gradient matrix $\mathbf{V} = \nabla u$ of the user's (time-varying) utility function and then project back to the problem's feasible region when the user's power constraints are violated. To that end, some straightforward matrix calculus yields:

$$\mathbf{V} = \nabla u = \frac{P_{\max}}{P_c + P_{\max}} \left[\mathbf{A} + \frac{\mathrm{tr}(\mathbf{A}\mathbf{Q}) - R(\mathbf{Q})}{P_c} \cdot \mathbf{I}\right], \tag{22}$$

where $\mathbf{Q}$ is given by (18) and

$$\mathbf{A} \equiv \nabla R(\mathbf{Q}) = \widetilde{\mathbf{H}}^\dagger [\mathbf{I} + \widetilde{\mathbf{H}} \mathbf{Q} \widetilde{\mathbf{H}}^\dagger]^{-1} \widetilde{\mathbf{H}}. \tag{23}$$

The above expression shows that $\mathbf{V}$ can be calculated at the transmitter as a function of the connection's effective channel matrix $\widetilde{\mathbf{H}}$ (which, in turn, can be estimated at the receiver and then fed back to the transmitter via a dedicated radio channel or as part of TDD downlink subframe). Moreover, since $\mathbf{V}$ is a bounded function of $\widetilde{\mathbf{H}}$ and the channel matrices $\widetilde{\mathbf{H}}(n)$ are assumed bounded, the induced sequence of gradient matrices $\mathbf{V}(n) \equiv \nabla u(\mathbf{X}(n); n)$ will also be bounded. We will therefore assume that there exists a constant $V_0 > 0$ such that

$$\|\mathbf{V}(n)\| \leq V_0 \quad \text{for all } n = 1, 2, \dots, \tag{24}$$

where $\|\mathbf{V}\| = \mathrm{tr}(\mathbf{V}^\dagger \mathbf{V})^{1/2}$ denotes the (Frobenius) matrix norm of $\mathbf{V}$.

[6] See also [12] for a similar use of the Charnes–Cooper transformation in the context of energy efficiency maximization.

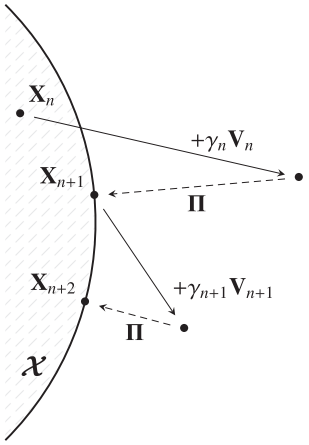


Fig. 1. Schematic representation of the recursive learning scheme (OGA).

In view of the above, and assuming for the moment perfect knowledge of $\mathbf{V}(n)$ at the transmitter, we will consider the matrix-valued online gradient ascent scheme

$$\mathbf{X}(n+1) = \mathbf{\Pi}(\mathbf{X}(n) + \gamma_n \mathbf{V}(n)), \tag{OGA}$$

where $\gamma_n > 0$ is a nonincreasing step-size sequence and $\mathbf{\Pi}$ denotes the matrix projection map:

$$\mathbf{\Pi}(\mathbf{Y}) = \arg\min_{\mathbf{X}\in\mathcal{X}} \|\mathbf{X} - \mathbf{Y}\|^2. \tag{25}$$

As we show in Appendix A, the matrix projection $\mathbf{\Pi}(\mathbf{Y})$ can be calculated by the simple expression:

$$\mathbf{\Pi}(\mathbf{Y}) = \mathbf{U} \cdot \operatorname{diag}(\boldsymbol{\pi}(\mathbf{y})) \cdot \mathbf{U}^\dagger, \tag{26}$$

where the tuple $(\mathbf{y}, \mathbf{U})$ diagonalizes $\mathbf{Y}$ (i.e. $\mathbf{Y} = \mathbf{U} \cdot \operatorname{diag}(\mathbf{y}) \cdot \mathbf{U}^\dagger$) and

$$\pi_i(\mathbf{y}) = \begin{cases} 0 & \text{if } y_i < 0, \\ y_i & \text{if } y_i \geq 0 \text{ and } \sum_j [y_j]_+ < 1, \\ [y_i - \lambda]_+ & \text{if } y_i \geq 0 \text{ and } \sum_j [y_j]_+ \geq 1, \end{cases} \tag{27}$$

with $\lambda > 0$ chosen so that $\sum_{i:y_i\geq 0} [y_i - \lambda]_+ = 1$.[7]

The iterative process (OGA) will be the main focus of our paper, so we proceed with some remarks (see also Fig. 1 for a schematic representation and Alg. 1 for a pseudocode implementation).

a) Properties: From a practical point of view, (OGA) has the following desirable attributes:

(P1) *Distributedness:* users require the same information as in distributed water-filling [49], [50].

(P2) *Statelessness:* users do not need to know the state of the system (e.g. the number of users in the network or its topology).

(P3) *Reinforcement:* users tend to become more energy-efficient based on their past observations.

[7]Recall here that $\mathbf{Y}(n)$ is Hermitian (because $\mathbf{V}(n)$ is Hermitian for all n), so its eigenvalues are real. Just as in water-filling methods [49], the Lagrange multiplier $\lambda > 0$ can then be calculated by sorting $\mathbf{y}$ and performing a line search for λ.

Algorithm 1. Online gradient ascent (OGA) for dynamic energy efficiency maximization

Parameter: variable step-size sequence $\gamma_n > 0$.
Initialize: $n \leftarrow 0$; $\mathbf{X} \leftarrow 0$.
Repeat
  $n \leftarrow n + 1$;
  {*Pre-transmission phase: set signal covariance matrix*}
  $\mathbf{Q} \leftarrow P_c P_{\max}/(P_c + P_{\max}(1 - \operatorname{tr}\mathbf{X})) \cdot \mathbf{X}$;
  transmit;
  {*Post-transmission phase: receive feedback and update*}
  get $\widetilde{\mathbf{H}}$;
  $\mathbf{A} \leftarrow \widetilde{\mathbf{H}}^\dagger \left[\mathbf{I} + \widetilde{\mathbf{H}}\mathbf{Q}\widetilde{\mathbf{H}}^\dagger\right]^{-1} \widetilde{\mathbf{H}}$;
  $\mathbf{V} \leftarrow P_{\max}/(P_c + P_{\max}) \left(\mathbf{A} + \left[\operatorname{tr}(\mathbf{A}\mathbf{Q}) - R(\mathbf{Q})/P_c \cdot \mathbf{I}\right]\right.$;
  $\mathbf{X} \leftarrow \mathbf{\Pi}(\mathbf{X} + \gamma_n \mathbf{V})$;
until transmission ends.

(P4) *Asynchronicity:* there is no need for a global update timer or any signaling/coordination between users; in particular, any subset of users may be updating at any given instance.

Regarding the algorithm's required feedback, we should note that each transmitter has the same information requirements as in distributed water-filling [8], [31], [49], [50]. In particular, as pointed out in [8], [50], a given transmitter does not need to know the other users' channels or covariance matrices, except via the aggregate MUI covariance matrix $\mathbf{W}$ at the receiver. This matrix is simply the expectation of the aggregate signal $\mathbf{W} = \mathbb{E}[\mathbf{y}\mathbf{y}^\dagger]$ reaching the receiver, so it can calculate it and feed it back to its connected transmitters (e.g. by broadcasting or as part of the downlink TDD phase). As a result, to calculate $\widetilde{\mathbf{H}}$ and $\mathbf{V}$, a user only requires this aggregate (and, hence, stateless) information and the knowledge of his actual channel matrix $\mathbf{H}$ (which is then used to calculate the effective channel matrix $\widetilde{\mathbf{H}}$).

b) Computational complexity: From a computational standpoint (which is crucial in massive MIMO systems), each iteration of Algorithm 1 requires a number of elementary binary operations which is polynomial (with a low degree) on the number of transmit/receive antennas and the number of subcarriers. Specifically, letting $S = \max\{M, N\}$ and recalling that $\widetilde{\mathbf{H}}$ and $\mathbf{Q}$ consist of K diagonal blocks, the required matrix multiplication and inversion steps for $\mathbf{A}$ and $\mathbf{V}$ carry a complexity of $\mathcal{O}(KS^\omega)$, with the complexity exponent ω being as low as 2.373 if fast Coppersmith–Winograd multiplication methods are employed [51]. As for the projection step $\mathbf{X} = \mathbf{\Pi}(\mathbf{Y})$, Eqs. (26) and (27) show that it can also be carried out in $\mathcal{O}(KS^\omega)$ operations: the diagonalization in (26) involves $\mathcal{O}(KS^\omega)$ steps while (27) only requires $\mathcal{O}(KM)$ operations for calculating the projection to the simplex [52].

With all this in mind, our main result for (OGA) is as follows:

Theorem 1: Assume that (OGA) is run with a variable step-size γ_n such that $\gamma_n \to 0$ and $n\gamma_n \to \infty$. Then, the induced transmit policy $\mathbf{Q}(n)$ leads to no regret in the online energy efficiency maximization problem (OEE). Specifically, (OGA) enjoys the cumulative regret bound

$$\operatorname{Reg}(T) \leq \frac{1}{\gamma_T} + \frac{1}{2} V_0^2 \sum_{n=1}^{T} \gamma_n, \tag{28}$$

or, using a step-size sequence of the form $\gamma_n = \gamma n^{-1/2}$:

$$\mathrm{Reg}(T) \leq \frac{1+\gamma^2 V_0^2}{\gamma}\sqrt{T}. \tag{29}$$

Proof: See Appendix B. ■

The anytime bound (28) will be our core performance guarantee, so some remarks are in order:

a) Fine-tuning γ_n: Theorem 1 shows that taking $\gamma_n \propto n^{-\alpha}$ for some $\alpha \in (0,1)$ leads to a regret guarantee that is $\mathcal{O}(T^\omega)$ with $\omega = \max\{\alpha, 1-\alpha\}$. [8] As such, (29) captures the optimal asymptotic behavior of the bound (28) for step-size sequences of the form $\gamma_n = \gamma/n^\alpha$. In fact, if V_0 can be estimated by the transmitter beforehand, the step-size parameter γ can be fine-tuned further so as to minimize the coefficient of $T^{1/2}$ in (29). Doing just that gives $\gamma = 1/V_0$, and provides the optimized bound:

$$\mathrm{Reg}(T) \leq 2V_0\sqrt{T}. \tag{30}$$

By exploiting equations (22), (23), and (24), we obtain the following expression for V_0 in terms of the system parameters:

$$V_0 = \frac{P_{\max}}{P_c + P_{\max}}\left(\sqrt{KN} + \frac{P_{\max}}{P_c}\left(\sqrt{KN}+1\right)\sqrt{KM}\right)\|\widetilde{\mathbf{H}}\|^2. \tag{31}$$

This expression reveals that the guarantee (30) is $\mathcal{O}(K\sqrt{MN})$, so the algorithm's overall regret is at most linear in the number of subcarriers and the number of transmit/receive antennas (assuming that M and N are of the same order). This guarantee is key for massive MIMO systems (where the number of transmit/receive antennas can grow fairly large) because it provides a worst-case estimate for the system's convergence. That said, (30) only becomes tight in adversarial environments (e.g. in the presence of jamming); in typical scenarios, the user's regret usually decays much faster and the system attains a stable, no-regret state within a few iterations, even for large numbers of antennas per user – cf. the detailed discussion in Sec. IV.

b) The static case: If the user's effective channels remain constant over time and $\mathbf{Q}(n)$ is a no-regret policy, a straightforward concavity argument can be used to show that $\max_n \mathrm{EE}(\mathbf{Q}(n))$ converges to the solution of the (static) EE maximization problem (9), [24], [53]. In this way, (OGA) can also be seen as a provably convergent low-cost algorithm for solving (9); furthermore, as we show in what follows, this convergence result continues to hold even in the presence of imperfect CSIT and measurement errors.

c) Initialization: The agnostic initialization $\mathbf{X}(0) = 0$ of Algorithm 1 means that the focal transmitter remains effectively silent during the first transmission frame (recall that $\mathbf{Q} \propto \mathbf{X}$). As such, the first iteration of (OGA) can be seen as a "handshake" that allows the transmitter to estimate his effective wireless medium before starting the actual transmission of data frames. If the transmitter begins with a given belief regarding his effective channel conditions, the algorithm can be initialized more aggressively in a manner consistent with the user's initial expectations (setting for instance $\mathbf{X} = (KM)^{-1}\mathbf{I}$ for uniform power allocation across subcarriers and antennas). In so doing, the regret bound (28) can be tightened further but this only makes a significant difference if the transmission horizon T is very short. Indeed, the influence of the initialization step is overtaken after only a few frames by the other users' behavior (which changes the effective channel) and becomes negligible as T grows large.

d) Logarithmic regret under fair channel conditions: As stated, Theorem 1 provides a worst-case guarantee which holds without any further caveats on the evolution of the channels from one stage to the next (other than basic information-theoretic hypotheses that allow the receiver to decode the transmitter's signal). As such, another important question that arises is whether we can achieve stronger performance guarantees under the additional hypothesis that channel conditions do not become too bad.

To quantify this, note first that the Shannon rate function $R(\mathbf{Q}) = \log\det(\mathbf{I} + \widetilde{\mathbf{H}}\mathbf{Q}\widetilde{\mathbf{H}}^\dagger)$ is strongly concave in $\mathbf{Q}$, and its strong concavity constant is an increasing function of the singular values of $\widetilde{\mathbf{H}}$.[9] Accordingly, since the user's energy efficiency function $\mathrm{EE}(\mathbf{Q})$ can be expressed as a perspective transformation $R(\mathbf{Q}) \mapsto xR(\mathbf{X}/x)$, the same will also hold for the strong concavity constant of $u(\mathbf{X})$ over $\mathcal{X}$ [33], [48]. Thus, following [54], if we assume that

$$\mathrm{Hess}(u(\mathbf{X}; n)) \leq -a\,\mathbf{I}, \tag{32}$$

for some $a > 0$ and for all $n = 1, 2, \ldots$, $\mathbf{X} \in \mathcal{X}$, we obtain the following stronger result:

Proposition 1: Assume that (OGA) is run with the step-size sequence $\gamma_n = \gamma/n$ for some $\gamma \geq a^{-1}$. Then, the induced transmit policy $\mathbf{Q}(n)$ enjoys the logarithmic regret bound:

$$\mathrm{Reg}(T) \leq \frac{1}{2}\gamma V_0^2(1+\log T). \tag{33}$$

Proof: See Appendix B. ■

Proposition 1 provides us with an important rule of thumb for choosing the step-size sequence of Algorithm 1. On the one hand, if the user expects that his effective channel can become arbitrarily bad (e.g. due to network congestion or deep fading events), the optimized bound (30) shows that the OGA algorithm should be run with a $n^{-1/2}$ step-size that allows higher adaptability to strongly varying channels. Otherwise, if the user expects reasonable channel quality over the entire transmission horizon, the "softer" step-size $\gamma_n \propto n^{-1}$ minimizes the danger of overcompensating for suboptimal transmit directions and allows the user to converge to a no-regret state faster.

C. Learning Under Uncertainty

A key assumption in our analysis so far is that the transmitter has access to perfect CSI and MUI measurements with which to calculate the gradient matrices $\mathbf{V}(n)$ at each stage. In practice however, factors such as pilot contamination, sparse feedback and imperfect channel sampling could have a deleterious effect

[8] To see this, simply note that $\sum_{n=1}^{T} n^{-\alpha} = \mathcal{O}(T^{1-\alpha})$ for large T and $\alpha \in (0,1)$.

[9] Recall here that a function f is strongly concave with constant $c > 0$ if $\mathrm{Hess}(f) \preccurlyeq -c\mathbf{I}$.

on the algorithm's performance. Our goal in this section will thus be to analyze (OGA) in the presence of uncertainty and feedback imperfections.

To formalize this, we will assume that, at each stage $n = 1, 2, \ldots$ of the process, the transmitter observes a noisy estimate $\hat{\mathbf{V}}(n)$ of $\mathbf{V}(n)$ satisfying the statistical hypotheses:

(H1) *Unbiasedness:*

$$\mathbb{E}[\hat{\mathbf{V}}(n) \mid \mathbf{Q}(n-1)] = 0. \tag{H1}$$

(H2) *Tame error tails:*

$$\mathbb{P}(\|\hat{\mathbf{V}}(n) - \mathbf{V}(n)\| \geq z) \leq B/z^{\beta}, \tag{H2}$$

for some $B > 0$ and for some $\beta > 2$.

Both hypotheses are fairly mild from a practical point of view. First, the unbiasedness hypothesis (H1) simply amounts to asking that there is no *systematic* error in the user's measurements [8], [55]–[60]. Secondly, Hypothesis (H2) is a bare-bones assumption on the probability of observing very high errors [8], and it represents a significant relaxation of the common assumption used in communications and signal processing whereby errors are assumed to lie in a bounded uncertainty region [56]–[60]. In particular, Hypothesis (H2) is met by all common error distributions, including Gaussian/sub-Gaussian, uniform, exponential, and log-normal distributed errors [55]–[60].

Remarkably, under these minimal hypotheses, we have:

Theorem 2: Assume that (OGA) is run with noisy measurements $\hat{\mathbf{V}}(n)$ satisfying Hypotheses (H1) and (H2), and with a variable step-size sequence of the form $\gamma_n = \gamma/n^{\alpha}$ for some $\alpha \in (2/\beta, 1)$. Then, the induced transmit policy $\mathbf{Q}(n)$ leads to no regret (a.s.) and enjoys the mean regret bound

$$\mathbb{E}[\mathrm{Reg}(T)] \leq \frac{1}{\gamma_T} + \frac{\hat{V}_0^2}{2}\sum_{n=1}^{T}\gamma_n, \tag{34}$$

where $\hat{V}_0^2 = \sup_n \mathbb{E}[\|\hat{\mathbf{V}}(n)\|^2]$.

Theorem 2 (proven in Appendix C) will be our main result in the context of dynamic energy efficiency maximization under imperfect CSI. A a few remarks are thus in order:

a) Step-sizes vs. large error probabilities: The requirement $\alpha\beta > 2$ of Theorem 2 indicates a trade-off between the probability of observing very large errors and achieving low regret (34). Specifically, if the error distribution of $\mathbf{Z}(n) = \hat{\mathbf{V}}(n) - \mathbf{V}(n)$ has very heavy tails (i.e. (H2) does not hold for $\beta \gg 2$), Algorithm 1 must be bootstrapped with a step-size sequence $\gamma_n \propto 1/n^{\alpha}$ for some $\alpha \approx 1$. In so doing however, the first term of (34) becomes almost linear, so the user might experience relatively high regret on average (due to the high probability of observing very large errors). On the other hand, if the tails of $\hat{\mathbf{V}}(n)$ are lighter (for instance, the standard case of Gaussion errors exhibits, exponentially thin tails, so (H2) holds for all β), Algorithm 1 can be employed with a more adaptive step-size sequence that guarantees a lower regret bound.

In particular, if (H2) holds for some $\beta > 4$, (OGA) can be used with a step-size sequence of the form $\gamma_n = \gamma n^{-1/2}$ which achieves the optimal behavior of (34), viz.

$$\mathbb{E}[\mathrm{Reg}(T)] \leq \frac{1+\gamma^2\hat{V}_0^2}{\gamma}\sqrt{T}. \tag{35}$$

Thus, if the mean square bound $\hat{V}_0^2 = \sup_n \mathbb{E}[\|\hat{\mathbf{V}}(n)\|^2]$ can be estimated ahead of time,[10] the step-size sequence γ_n can be optimized further. More precisely, as in the deterministic case, the coefficient of $T^{1/2}$ in (35) is minimized when $\gamma = 1/\hat{V}_0$, so we obtain the optimized bound:

$$\mathbb{E}[\mathrm{Reg}(T)] \leq 2\hat{V}_0\sqrt{T}. \tag{36}$$

b) The estimation process: The no-regret properties of (OGA) under uncertainty rely on the availability of statistically unbiased measurements $\hat{\mathbf{V}}$ of $\mathbf{V}$. In turn, given that users have perfect knowledge of their individual transmit covariance matrices, this requirement boils down to constructing an unbiased estimator of the matrix $\mathbf{A} = \tilde{\mathbf{H}}^{\dagger}\left(\mathbf{I} + \tilde{\mathbf{H}}\mathbf{Q}\tilde{\mathbf{H}}^{\dagger}\right)^{-1}\tilde{\mathbf{H}}$. In our context, this can be accomplished via the statistical sampling process of [7], [8] which provides an unbiased estimator of $\mathbf{A}$ with exponentially decaying error tails (i.e. (H2) holds for all $\beta > 2$). However, due to space limitations we will not address this question in more detail here.

c) Fair channel conditions and noise: As before, the regret guarantee (34) can be tightened significantly if the user's effective channel conditions satisfy (32). In that case, if (OGA) is run with a variable step-size sequence of the form $\gamma_n \propto 1/n$, we obtain the following stochastic analogue of Proposition 1:

Proposition 2: With notation as in Theorem 2, assume that (OGA) is run with noisy measurements and a variable step-size sequence $\gamma_n = \gamma/n$ for some $\gamma \geq a^{-1}$ with a. Then, the induced transmit policy $\mathbf{Q}(n)$ leads to no regret (a.s.) and enjoys the mean guarantee:

$$\mathbb{E}[\mathrm{Reg}(T)] \leq \frac{1}{2}\gamma\hat{V}_0^2(1+\log T). \tag{37}$$

Proof: See Appendix C. ∎

As in the perfect CSI case, Proposition 2 provides a rule of thumb for achieving lower regret faster when the user's (effective) wireless medium is not too bad: as long as (32) holds for some $a > 0$, the user can achieve logarithmic regret, even with very noisy measurements.

IV. Numerical Results

In this section, we assess the performance of the OGA algorithm via numerical simulations. Even though we only present here a representative subset of these results, our conclusions apply to a wide range of wireless network parameters and specifications.

Our setup is as follows: we consider a cellular OFDMA wireless network occupying a 10 MHz band divided into 1024 subcarriers around a central frequency of $f_c = 2.5$ GHz. Wireless signal propagation is modeled following the well-known COST Hata model [40], [41] and the spectral noise density is taken to be -174 dBm/Hz at 20 °C (for a detailed overview of simulation parameters, see Table I). Network coverage is provided by 19 hexagonal cells (each with a radius of 1 km) that form a honeycomb pattern spanning an urban area with wireless user density $\rho = 500$ users/km^2. To minimize complexity, OFDM

[10] Note here that $\hat{V}_0$ is guaranteed to be finite on account of Hypothesis (H2).

TABLE I
WIRELESS NETWORK SIMULATION PARAMETERS

Parameter	Value
Time frame duration	5 ms
Central frequency	2.5 GHz
Total bandwidth	11.2 MHz
OFDM subcarriers	1024
Spectral noise density (20 °C)	−174 dBm/Hz
User speed	[3, 130] km/h
Maximum transmit power	$P_{\max} = 33$ dBm
Non-radiative power	$P_c = 20$ dBm
Transmit antennas per device	$M = 4$
Receive antennas per link	$N = 8$

subcarriers are allocated to wireless users in each cell following a simple randomized access scheme that assigns different users to disjoint subcarrier sets [61].

We focus on a set of $U = 15$ transmitting users that are located in different cells (following a Poisson point process sampling) and share $K = 8$ common subcarriers. Each wireless transmitter is further assumed to have $M = 4$ transmit antennas, a maximum transmit power of $P_{\max} = 40$ dBm and circuit (non-radiative) power consumption of $P_c = 20$ dBm. At the receiver end, we consider $N = 8$ receive antennas per connection and a receiver noise figure of 7 dB. Finally, communication occurs over a time-division duplexing (TDD) scheme with frame duration $T_f = 5$ ms: transmission takes place during the uplink (UL) subframe while the receivers process the received signal and provide feedback during the downlink (DL) subframe; upon reception of the feedback, the users update their power profiles following Alg. 1 and the process repeats.

A. Static Channels

For benchmarking purposes, our first simulation scenario studied stationary users with static channel conditions (so the variability of a user's effective channel matrix is only due to the modulation of the interfering users' transmit characteristics). Each user is assumed to run the OGA algorithm with a variable step-size of the form $\gamma_n \propto 1/\sqrt{n}$ and an agnostic initialization with initial transmit power $P_0 = P_{\max}/2 = 26$ dBm spread evenly across antennas and subcarriers. Our simulation results are presented in Fig. 2 where, to minimize graphical clutter, we only plot the relevant information for four users with diverse channel characteristics.

First, in Fig. 2(a), we plot the users' transmit power under OGA. As can be seen, even though users change their power by several dBm, the algorithm quickly converges after an initial transient phase. Similarly, in Fig. 2(b), we plot the users' transmit rate over all subcarriers (normalized by the bandwidth and thus measured in bps/Hz). We see here that users who reduce power by more than 10 dBm (Users 3 and 4) experience a commensurate drop in spectral efficiency (of the order of a few bps/Hz). On the other hand, users that decrease power only by a little achieve higher rates because the OGA algorithm leads to a more efficient allocation of power over subcarriers and antennas. Nonetheless, in all cases, we observe a dramatic increase in energy efficiency over the users' initial (uniform) power allocation policy, ranging from $\approx 200\%$ to more than 600%.[11]

[11] Contrary to Fig. 2(b), we do not normalize the users' energy efficiency by the bandwidth, so it is measured in Mb/J.

In fact, as we see in Fig. 2(c), after some slight oscillations during the first few iterations (the algorithm's transient phase), the system rapidly reaches a state where users have no incentive to change their individual power profiles (a *Nash equilibrium* in game-theoretic terms). This equilibration is consistent with the no-regret properties of the OGA algorithm: as predicted by Theorem 1 and shown in Fig. 2(d), the users' regret quickly decays to zero.

B. Time-Varying Channels and Mobility

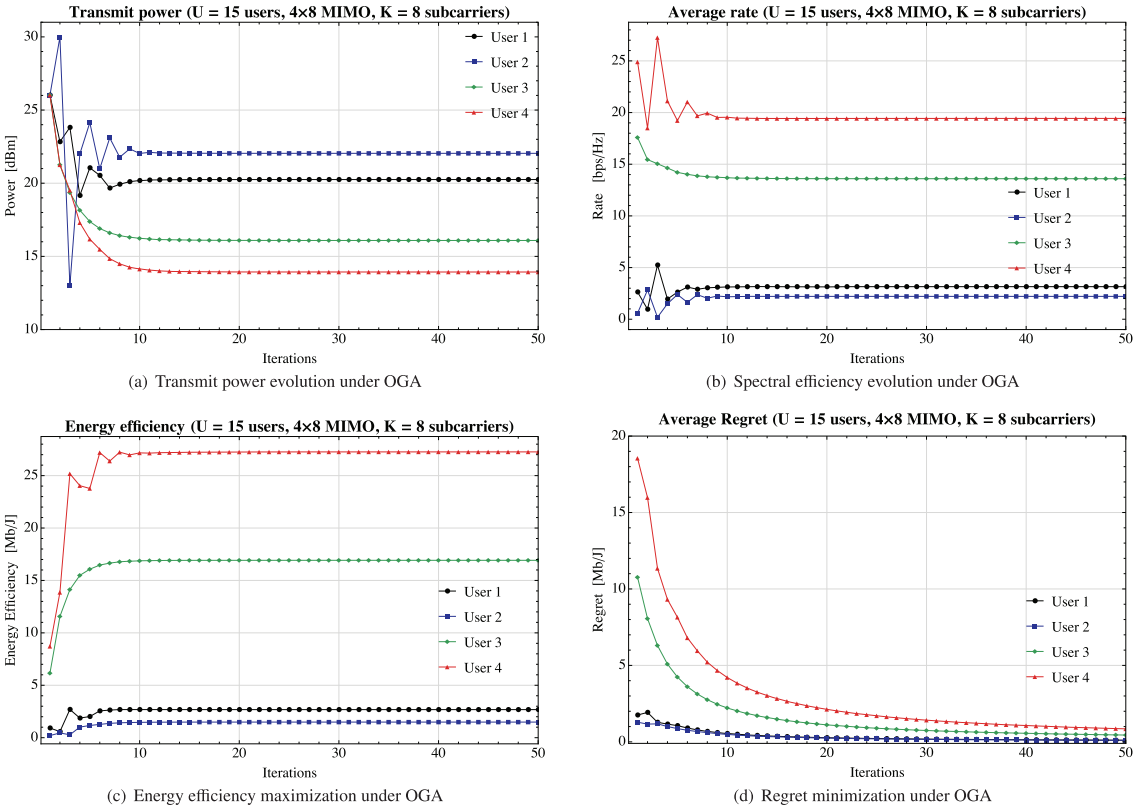
To account for dynamic network conditions, we also consider in Fig. 3 the case of mobile users whose channels vary with time due to Rayleigh fading, path loss fluctuations, etc. For simulation purposes, we used the extended typical urban (ETU) model for the users' environment and the extended pedestrian A (EPA) and extended vehicular A (EVA) models to simulate pedestrian (3–5 km/h) and vehicular (30–130 km/h) movement respectively [62]. For reference, the focal users' channel gains ($\mathrm{tr}(\mathbf{H}\mathbf{H}^\dagger)$) have been plotted in Fig. 3(a). Despite the channels' variability, Fig. 3(b) shows that the users attain a no-regret state in a few iterations, even under rapidly changing channel conditions (cf. the case of Users 2 and 4 with an average speed of 30 km/h and 130 km/h respectively).

For completeness, we also plot in Figs. 3(c) and 3(d) the achieved energy efficiency for a pedestrian and a vehicular user, and we compare it to a) the user's initial (uniform) power allocation policy; b) the (fixed) oracle solution of the offline maximization problem (12), calculated with full knowledge of the future; and c) the (dynamic) optimum policy $\mathbf{Q}^*(n) = \arg\max_{\mathbf{Q}} \mathrm{EE}(\mathbf{Q}; n)$ which represents the instantaneous optimum policy in terms of energy efficiency (and which is again calculated with perfect foresight and knowledge of the future). Remarkably, even under rapidly changing channel conditions, the users' achieved energy efficiency tracks its (evolving) maximum value remarkably well and consistently outperforms the fixed oracle solution (a fact which is consistent with the negative regret observed in Fig. 3(b)).

An intuitive explanation for the adaptability of OGA is provided by Figs. 3(e) and 3(f) where we plot the transmit power of the dynamic optimum policy, the OGA scheme and the fixed oracle solution for the same users as in Figs. 3(c) and 3(d). Even though the dynamic optimum policy $\mathbf{Q}^*(n)$ may change significantly from one frame to the next, $\mathrm{tr}(\mathbf{Q}^*(n))$ remains roughly constant (within a few dBm) over the entire transmission horizon. The OGA algorithm learns this power level in a few iterations and stays close to it throughout the transmission horizon. As a result, the users' achieved energy efficiency remains itself very close to its maximum value.

C. Robustness to Observation Noise and Scalability for Large Antenna Numbers

To assess the robustness of the OGA algorithm in the presence of observation noise and measurement errors, the simulation cycle above was repeated in Fig. 4 for the case where users only have access to noisy gradient observations as in Section III-C. Also, to study the algorithm's scalability in the massive MIMO regime (large number of antennas), we increased the number of transmit antennas to $M = 8$ and the



(a) Transmit power evolution under OGA

(b) Spectral efficiency evolution under OGA

(c) Energy efficiency maximization under OGA

(d) Regret minimization under OGA

Fig. 2. Performance of the Algorithm 1 under static channel conditions. The system converges within a few iterations to an equilibrium state (Fig. 2(c)) where users experience no regret (Fig. 2(d)).

number of receive antennas to $N = 128$; other than that, we used the same network simulation parameters as in Fig. 2. The intensity of the measurement noise was quantified via the relative error level of the estimator $\hat{\mathbf{V}}$, i.e. its standard deviation over its mean (so a relative error level of $\eta\%$ means that the observed matrix $\hat{\mathbf{V}}$ lies within $\eta\%$ of its true value). We then plotted the users' achieved energy efficiency under the OGA algorithm for noise levels $\eta = 20\%$ (moderate-to-high uncertainty) and $\eta = 100\%$ (very high uncertainty). As can be seen in Fig. 4, the system's rate of equilibration is adversely affected by the intensity of the noise; however, the system still equilibrates within a few tens of iterations and the users exhibit a drastic increase in energy efficiency (of the order of 150% and more), even in the presence of very high uncertainty.

Furthermore, to study the performance of the proposed algorithm in the massive MIMO regime, we also plotted in Fig. 5 the number of iterations required for the system to equilibrate for different numbers of antennas at the transmitter and the receiver end (M and N respectively). Importantly, as can be seen in Fig. 5, the OGA algorithm scales very well with the number of antennas and converges within a few iterations, even in the massive MIMO regime. Specifically, the algorithm requires around 10 iterations to converge for 8×92 MIMO systems and around 20 iterations for 16×128 systems.

D. The Cooperative Framework

As is well known in game theory, there is often a gap between unilaterally optimal solutions (such as Nash equilibria) and socially optimal ones (where the players' aggregate utility is maximized). As such, an important question that arises is the following: *what is the gap (if any) between the unilateral framework presented above and a cooperative setting where users can coordinate their actions in order to maximize the system's overall energy efficiency?*

To formalize this question, we consider below a system consisting of U wireless users that transmit to a common receiver (for simplicity, we assume that all users have similar circuit and maximum power characteristics as in the rest of this section).In this case, the system's achievable sum rate [37], [49] is given by:

$$R_{\text{sys}}(\mathbf{Q}_1, \ldots, \mathbf{Q}_U) = \log\det\left(\mathbf{I} + \textstyle\sum_u \mathbf{H}_u \mathbf{Q}_u \mathbf{H}_u^\dagger\right), \qquad (38)$$

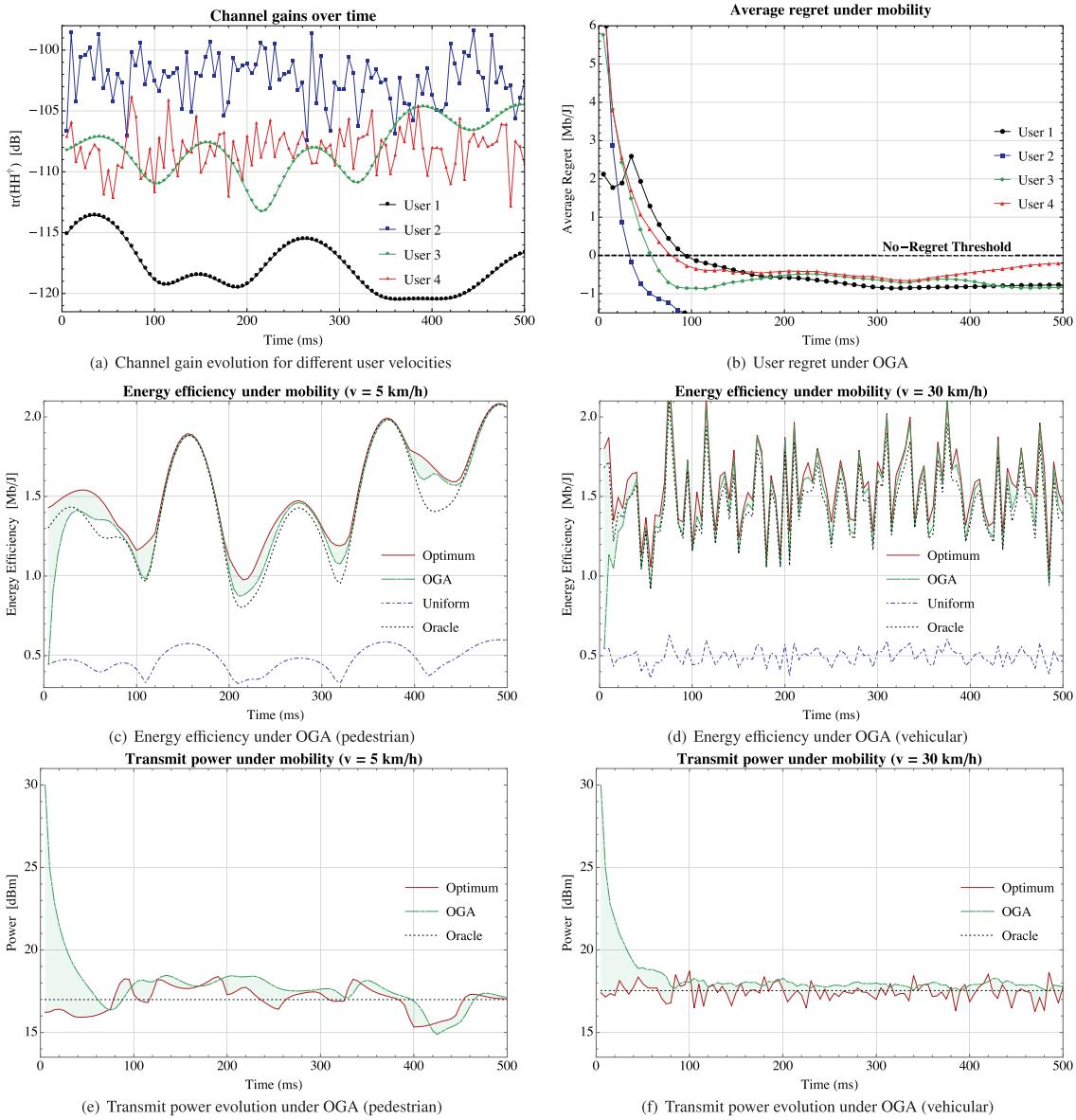
where, in obvious notation, $\mathbf{Q}_u$ denotes the covariance matrix of the u-th user. Accordingly, the *system's* energy efficiency will be given by the ratio

$$\text{EE}_{\text{sys}}(\mathbf{Q}_1, \ldots, \mathbf{Q}_U) = \frac{R_{\text{sys}}(\mathbf{Q}_1, \ldots, \mathbf{Q}_U)}{\sum_u (P_c + \text{tr}(\mathbf{Q}_u))}, \qquad (39)$$

leading to the maximization problem:

$$\begin{aligned} &\text{maximize} && \text{EE}_{\text{sys}}(\mathbf{Q}_1, \ldots, \mathbf{Q}_U), \\ &\text{subject to} && \mathbf{Q}_u \in \mathcal{Q},\ u \in \mathcal{U}. \end{aligned} \qquad (40)$$

To solve this problem, we can use the analysis of Section III to derive a cooperative gradient ascent (CGA) scheme that has



(a) Channel gain evolution for different user velocities

(b) User regret under OGA

(c) Energy efficiency under OGA (pedestrian)

(d) Energy efficiency under OGA (vehicular)

(e) Transmit power evolution under OGA (pedestrian)

(f) Transmit power evolution under OGA (vehicular)

Fig. 3. Performance of the OGA algorithm in a dynamic setting with mobile users moving at $v \in \{3, 30, 5, 130\}$ km/h. The users' achieved energy efficiency tracks its (evolving) maximum value remarkably well, even under rapidly changing channel conditions.

the same update structure as the unilateral OGA method but is instead adapted to the system-oriented problem (40) above. Because the details of this construction would take us too far afield, we omit them;[12] we only note that the resulting CGA method is no longer distributed.

Thanks to this cooperative method, we can assess the performance of the unilateral OGA algorithm in a cooperative framework and examine the gap between the two settings. Specifically, in Fig. 6, we plot the system's overall energy efficiency for $U = 5$ users that employ the CGA and OGA schemes (in the cooperative and unilateral frameworks respectively). In the algorithm's end state, the system's energy efficiency when users do not cooperate lies fairly close to that of the cooperative framework (within 5%). On the other hand, cooperative users achieve substantially higher values of system-wide energy efficiency in the transient regime – but at the cost of signaling overhead between the users.

Remark 4: The (Gaussian) multiple access channel (MAC) model described above is a special case of the interference channel in which the receivers are colocated and cooperative. As opposed to the general interference channel (where the centralized sum-capacity and capacity region are

[12] In brief, the CGA algorithm is constructed in three steps. We first use the Charnes–Cooper transformation to define a new set of variables $\mathbf{Z} = \mathbf{Z}(\mathbf{Q})$ that map the problem (40) to an equivalent concave problem with objective function $u_{\text{sys}}(\mathbf{Z})$ and feasible region $\mathcal{Z}$. The CGA recursion is then defined by replacing the gradient in (OGA) with $\nabla_{\mathbf{Z}} u_{\text{sys}}(\mathbf{Z})$ and then taking the closest point projection to $\mathcal{Z}$.

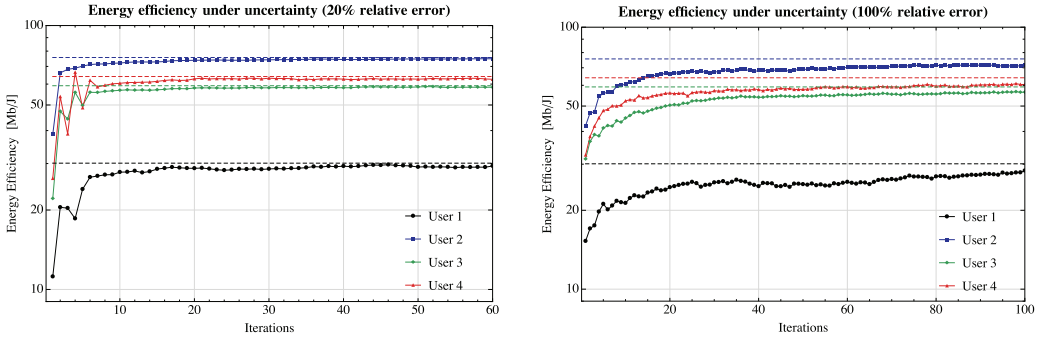


Fig. 4. Performance of Algorithm 1 with imperfect measurements and observation errors. Remarkably, even under very high uncertainty, the system converges within a few tens of iterations to a stable state (dashed lines) where users experience no regret.

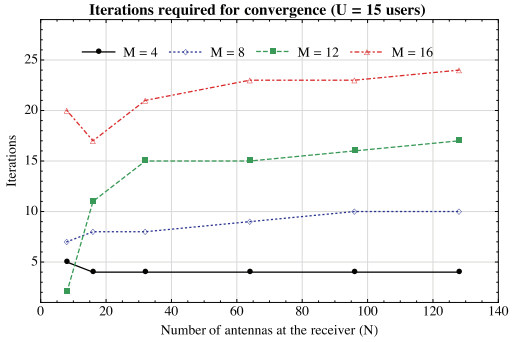


Fig. 5. The convergence speed of Algorithm 1 for different numbers of antennas at the transmitter and the receiver.

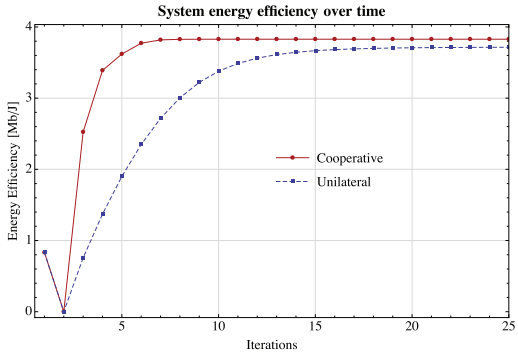


Fig. 6. Energy efficiency comparison between the unilateral and cooperative frameworks.

not known), the capacity region of the MAC is well understood and its sum-capacity is a concave function of the users' input signal covariance matrices. On the other hand, due to non-convexity issues that arise in the full interference channel, even the maximization of the users' achievable sum-rate under SUD is a very hard nonlinear problem in the general case. Therefore, when the figure of merit is the system's energy efficiency, the Charnes–Cooper transform (or other fractional programming methods) cannot be applied to compute the system's globally optimum transmit profile. In the recent paper [63], the authors attempt to circumvent this difficulty by taking a convexified relaxation (i.e. a lower bound estimate) of this problem and propose an algorithm to achieve a solution thereof. However, because it is hard to quantify the performance gap between this relaxed solution and the problem's global optimum, the methodology of [63] does not readily yield an accurate comparison between the cooperative and non-cooperative regime for the full interference channel (especially beyond the case of static channels). A comprehensive analysis of these issues lies beyond the scope of our paper so we focus on the MAC case where convexity allows us to derive below a cooperative learning algorithm with guaranteed convergence/no-regret properties.

V. Conclusions and Perspectives

In the context of multi-carrier MIMO systems where numerous interfering mobile users co-exist, the temporal variability of the system cannot be ignored when targeting highly energy-efficient communications. To tackle this issue, we introduced an online optimization framework for the study of energy efficiency maximization in dynamically varying networks, and we proposed an adaptive transmit policy that allows users to attain a "no-regret" state. Importantly, the proposed policy is distributed, asynchronous, computationally simple, and it only requires minimal, strictly causal and (possibly) noisy feedback. Specifically, under very mild assumptions on the statistics of the error process, we showed that the users' average regret after T epochs decays as $\mathcal{O}(T^{-1/2})$, a bound which is further improved to $\mathcal{O}(\log T/T)$ if the users do not experience arbitrarily bad channel conditions. As a result, users are able to track their most efficient transmit power profile with modest feedback requirements, even under rapidly changing channel conditions (corresponding to highly mobile users): indeed, our simulations show that users could gain up to 600% in energy efficiency over fixed/uniform power allocation policies in realistic network environments.

An important theoretical question which arises is whether the system converges to an equilibrium state if all users employ a no-regret policy (our numerical simulations show that this indeed the case over a wide region of system parameters). Additionally, different throughput-per-power models accounting for the probability of outage can also be considered and would require a modification of the proposed transmit policy. We intend to explore these directions in future work.

APPENDIX
TECHNICAL PROOFS

Throughout this appendix (and unless explicitly mentioned otherwise), all matrices are assumed Hermitian and of dimension $D = KM$. Additionally, the stage variable n will be written as a subscript instead of as an argument – i.e. we will write $\mathbf{X}_n$ and $u_n(\mathbf{X})$ instead of $\mathbf{X}(n)$ and $u(\mathbf{X}; n)$ respectively. We do so in order to reduce the notational clutter caused by an overflow of parentheses; since we will not require a subcarrier index, there is no fear of ambiguity.

A. Matrix Projections

We first prove that the projection map $\mathbf{\Pi}(\mathbf{Y})$ is given by the explicit formula (26). To that end, simply note that $\mathbf{\Pi}(\mathbf{Y})$ can be expressed equivalently as the solution to the maximization problem

$$\begin{aligned} &\text{maximize} && \operatorname{tr}(\mathbf{YX}) - \tfrac{1}{2}\|\mathbf{X}\|^2, \\ &\text{subject to} && \mathbf{X} \in \mathcal{X}. \end{aligned} \tag{41}$$

However, if $\mathbf{Y} = \mathbf{U}\Delta\mathbf{U}^\dagger$ is a diagonalization of $\mathbf{Y}$, the objective of (41) can be written as:

$$\operatorname{tr}(\mathbf{YX}) - \tfrac{1}{2}\|\mathbf{X}\|^2 = \operatorname{tr}(\Delta\mathbf{U}^\dagger\mathbf{XU}) - \tfrac{1}{2}\operatorname{tr}(\mathbf{U}^\dagger\mathbf{XUU}^\dagger\mathbf{XU}). \tag{42}$$

Thus, given that $\mathbf{X} \in \mathcal{X}$ if and only if $\mathbf{UXU}^\dagger \in \mathcal{X}$, we readily get $\mathbf{\Pi}(\mathbf{Y}) = \mathbf{U}\mathbf{\Pi}(\Delta)\mathbf{U}^\dagger$, so it suffices to solve (41) for diagonal $\mathbf{Y}$.

We first show that $\mathbf{\Pi}(\mathbf{Y})$ is itself diagonal if $\mathbf{Y} = \operatorname{diag}(\mathbf{y})$ for some $\mathbf{y} \in \mathbb{R}^D$. Indeed, we have:

$$\operatorname{tr}(\mathbf{YX}) - \tfrac{1}{2}\|\mathbf{X}\|^2 = \textstyle\sum_i y_i X_{ii}^2 - \frac{1}{2}\sum_{i,j}|X_{ij}|^2 \leq \sum_i (y_i X_{ii}^2 - \tfrac{1}{2}X_{ii}^2), \tag{43}$$

with equality if and only if $\mathbf{X}$ is diagonal. As a result, if $\mathbf{X}$ is a solution of (41), the diagonal matrix $\mathbf{X}'$ which coincides with $\mathbf{X}$ on the diagonal and has zero entries otherwise will also be a solution of (41); since (41) admits a unique solution, we conclude that $\mathbf{\Pi}(\mathbf{Y})$ must also be diagonal, as claimed. We are thus left to solve the maximization problem

$$\begin{aligned} &\text{maximize} && \textstyle\sum_j y_j x_j - \frac{1}{2}\sum_j x_j^2, \\ &\text{subject to} && x_j \geq 0, \quad \textstyle\sum_j x_j \leq 1. \end{aligned} \tag{44}$$

Writing $\lambda_j \geq 0$ and $\lambda \geq 0$ for the Lagrange multipliers of the constraints $x_j \geq 0$ and $\sum_j x_j \leq 1$ respectively, the first-order Karush–Kuhn–Tucker (KKT) conditions for (44) become:

$$y_j = x_j + \lambda - \lambda_j, \tag{45a}$$

$$\lambda_j x_j = 0, \quad \lambda(1 - \textstyle\sum_j x_j) = 0. \tag{45b}$$

Thus, to obtain the first branch of (27), simply note that if $y_j \leq 0$ but $x_j > 0$, we will also have $\lambda_j = 0$, so (45a) gives $y_j = x_j + \lambda > 0$, a contradiction. Likewise, if $\sum_{j:y_j\geq 0} y_j \leq 1$, setting $\lambda_j = \lambda = 0$ and $x_j = y_j$ for all j such that $y_j \geq 0$ is obviously a solution of (45), so we obtain the second branch of (27). Finally, to obtain the third branch of (27), note first that $\sum_{j:y_j\geq 0} x_j = 1$ if $\sum_{j:y_j\geq 0} y_j \geq 1$; otherwise, we would have $\lambda = 0$ and (45a) would give $y_j = x_j - \lambda_j \leq x_j$ whenever $y_j \geq 0$, implying in turn that $\sum_{j:y_j\geq 0} y_j \leq \sum_{j:y_j\geq 0} x_j < 1$, a contradiction. Accordingly, we are left to project the vector with components $y_i^+ = \big[y_i\big]_+$ to the unit simplex $\Delta = \{\mathbf{x} \in \mathbb{R}^D : x_j \geq 0 \text{ and } \sum_j x_j = 1\}$; this projection simply gives $x_i = [y_i^+ - \lambda]_+$ with $\lambda \geq 0$ such that $\sum_i [y_i^+ - \lambda]_+ = 1$ [52], so (26) follows.

B. No Regret With Perfect Feedback

The key step in bounding the user's regret is the inequality:

$$\operatorname{EE}(\mathbf{Q}^*) - \operatorname{EE}(\mathbf{Q}) = u(\mathbf{X}^*) - u(\mathbf{X}) \leq \operatorname{tr}[\mathbf{V}\cdot(\mathbf{X}^* - \mathbf{X})], \tag{46}$$

itself a simple consequence of the fact that $u(\mathbf{X})$ is concave in $\mathbf{X}$. Our proof follows the methodology of [35] where OGA methods where used in a vector (as opposed to matrix) setting with the special step-size sequence $\gamma_n \propto n^{-1/2}$ (as opposed to general γ_n). To be precise, we will establish the no-regret properties of (OGA) by showing that $\sum_{n=1}^T \operatorname{tr}[\mathbf{V}_n \cdot (\mathbf{X}^* - \mathbf{X}_n)] = o(T)$ for all $\mathbf{X}^* \in \mathcal{X}$ and for every matrix sequence $\mathbf{V}_n$.

Proof of Theorem 1: Letting $D_n = \frac{1}{2}\|\mathbf{X}^* - \mathbf{X}_n\|^2$, we get:

$$D_{n+1} = \frac{1}{2}\left\|\mathbf{X}^* - \mathbf{\Pi}(\mathbf{X}_n + \gamma_n\mathbf{V}_n)\right\|^2 \leq \frac{1}{2}\left\|\mathbf{X}^* - \mathbf{X}_n - \gamma_n\mathbf{V}_n\right\|^2, \tag{47}$$

on account of the definition of $\mathbf{\Pi}(\mathbf{Y})$ as the closest point to $\mathbf{Y}$ on $\mathcal{X}$. In this way, (47) yields:

$$D_{n+1} \leq D_n - \gamma_n \operatorname{tr}\big[\mathbf{V}_n \cdot (\mathbf{X}^* - \mathbf{X}_n)\big] + \frac{1}{2}\gamma_n^2\|\mathbf{V}_n\|^2, \tag{48}$$

and hence, after rearranging and summing over n, we obtain:

$$\begin{aligned} \sum_{n=1}^T \operatorname{tr}\big[\mathbf{V}_n \cdot (\mathbf{X}^* - \mathbf{X}_n)\big] &\leq \sum_{n=1}^T \gamma_n^{-1}(D_n - D_{n+1}) + \frac{1}{2}\sum_{n=1}^T \gamma_n\|\mathbf{V}_n\|^2 \\ &\leq \gamma_1^{-1}D_1 + \sum_{n=2}^T \left(\gamma_n^{-1} - \gamma_{n-1}^{-1}\right) D_n + \frac{1}{2}\sum_{n=1}^T \gamma_n\|\mathbf{V}_n\|^2 \\ &\leq \gamma_1^{-1}\Omega + \sum_{n=2}^T \left(\gamma_n^{-1} - \gamma_{n-1}^{-1}\right)\Omega + \frac{V_0^2}{2}\sum_{n=1}^T \gamma_n = \frac{1}{\gamma_T} + \frac{V_0^2}{2}\sum_{n=1}^T \gamma_n, \end{aligned} \tag{49}$$

where $\Omega \equiv \frac{1}{2}\max_{\mathbf{X},\mathbf{X}'\in\mathcal{X}}\|\mathbf{X} - \mathbf{X}'\|^2 = 1$. The fact that $\mathbf{X}_n$ leads to no regret then follows by noting that $1/(T\gamma_T) \to 0$ (by assumption) and that $T^{-1}\sum_{n=1}^T \gamma_n \to 0$ (since $\gamma_n \to 0$). ∎

Proof of Proposition 1: Reasoning as in [33], [54], the strong concavity assumption (32) gives:

$$u_n(\mathbf{X}^*) - u_n(\mathbf{X}_n) \leq \operatorname{tr}\big[\mathbf{V}_n(\mathbf{X}^* - \mathbf{X}_n)\big] - \frac{1}{2\gamma}\|\mathbf{X}^* - \mathbf{X}_n\|^2, \tag{50}$$

where we have used the fact that $a \geq \gamma^{-1}$. Thus, by summing over n and using (49), we obtain:

$$\begin{aligned} \operatorname{Reg}(T) &\leq \frac{1}{2}\sum_{n=2}^T \left(\gamma_n^{-1} - \gamma_{n-1}^{-1} - \gamma^{-1}\right)\|\mathbf{X}^* - \mathbf{X}_n\|^2 + \frac{1}{2}\sum_{n=1}^T \gamma_n\|\mathbf{V}_n\|^2 \\ &\leq \frac{1}{2}\gamma V_0^2\sum_{n=1}^T n^{-1} \leq \frac{1}{2}\gamma V_0^2(1 + \log T), \end{aligned} \tag{51}$$

where, in the second line, we used the fact that $\gamma_n^{-1} - \gamma_{n-1}^{-1} = n\gamma^{-1} - (n-1)\gamma^{-1} = \gamma^{-1}$. ∎

C. The Case of Imperfect Feedback

Proof of Theorem 2: As before, we begin with the basic inequality:

$$\mathrm{Reg}(T) \leq \max_{\mathbf{X}^* \in \mathcal{X}} \sum_{n=1}^{T} \mathrm{tr}\left[\mathbf{V}_n \cdot (\mathbf{X}^* - \mathbf{X}_n)\right], \quad (52)$$

with $\mathbf{X}_n$ defined via the (stochastic) recursion:

$$\mathbf{X}_{n+1} = \mathbf{\Pi}(\mathbf{X}_n + \gamma_n \hat{\mathbf{V}}_n). \quad (53)$$

Thus, for the first part of the theorem, we need to show that:

$$\sum_{n=1}^{T} \mathrm{tr}[\hat{\mathbf{V}}_n \cdot (\mathbf{X}^* - \mathbf{X}_n)] + \sum_{n=1}^{T} \mathrm{tr}\left[\mathbf{Z}_n \cdot (\mathbf{X}_n - \mathbf{X}^*)\right] = o(T) \quad \text{(a.s.)}, \quad (54)$$

where $\mathbf{Z}_n = \hat{\mathbf{V}}_n - \mathbf{V}_n$. The first term of (54) then becomes:

$$\sum_{n=1}^{T} \mathrm{tr}[\hat{\mathbf{V}}_n \cdot (\mathbf{X}^* - \mathbf{X}_n)] \leq \frac{\Omega}{\gamma_1} + \sum_{n=2}^{T} \left(\gamma_n^{-1} - \gamma_{n-1}^{-1}\right) D_n$$
$$+ \sum_{n=1}^{T} \frac{\gamma_n \|\hat{\mathbf{V}}_n\|^2}{2} \leq \frac{1}{\gamma_T} + \frac{V_0^2}{2} \sum_{n=1}^{T} \gamma_n + \mathcal{O}\left(\sum_{n=1}^{T} \gamma_n \|\mathbf{Z}_n\|^2\right), \quad (55)$$

where, as before, $\Omega \equiv \frac{1}{2} \max_{\mathbf{X}, \mathbf{X}' \in \mathcal{X}} \left\|\mathbf{X} - \mathbf{X}'\right\|^2 = 1$. We now claim that $\lim_{T\to\infty} T^{-1} \sum_{n=1}^{T} \gamma_n \|\mathbf{Z}_n\|^2 \to 0$ (a.s.). Indeed, let $z_n = \|\mathbf{Z}_n\|$ and choose $\varepsilon > 0$ such that $4\varepsilon \leq \alpha\beta - 2$ (recall that $\alpha\beta > 2$); Hypothesis (H2) then implies that $\mathbb{P}(z_n \geq n^{\alpha/2-\varepsilon/\beta}) \leq B/n^{\alpha\beta/2-\varepsilon} \leq B/n^{1+\varepsilon}$ for all n, so we obtain:

$$\sum_{n=1}^{\infty} \mathbb{P}\left(z_n \geq n^{\alpha/2-\varepsilon/\beta}\right) = \sum_{n=1}^{\infty} \mathcal{O}(1/n^{1+\varepsilon}) < \infty, \quad (56)$$

and hence, by the Borel-Cantelli lemma, we conclude that $\mathbb{P}(z_n \geq n^{\alpha/2-\varepsilon/\beta}$ for infinitely many $n) = 0$. In turn, this implies that $z_n^2 = \mathcal{O}\left(n^{\alpha-2\varepsilon/\beta}\right)$ almost surely, so we get (a.s.):

$$\sum_{n=1}^{T} \gamma_n \|\mathbf{Z}_n\|^2 = \mathcal{O}\left(\sum_{n=1}^{T} n^{-\alpha} n^{\alpha-2\varepsilon/\beta}\right) = \mathcal{O}\left(\sum_{n=1}^{T} 1/n^{2\varepsilon/\beta}\right) = o(T). \quad (57)$$

For the second term of (54), let $\xi_n = \mathrm{tr}\left[\mathbf{Z}_n \cdot (\mathbf{X}^* - \mathbf{X}_n)\right]$. Then, given that $\mathbf{X}_n$ is a deterministic function of $\mathbf{X}_{n-1}$ and $\hat{\mathbf{V}}_{n-1}$, we will also have $\mathbb{E}\left[\xi_n \mid \mathbf{X}_{n-1}\right] = 0$, i.e. ξ_n is a sequence of martingale differences. By the strong law of large numbers for martingale differences [64, Theorem 2.18], it then follows that $\lim_{T\to\infty} T^{-1} \sum_{n=1}^{T} \xi_n = 0$ (a.s.). As a result, combining this with (57), we get $\sum_{n=1}^{T} \mathrm{tr}\left[\hat{\mathbf{V}}_n \cdot (\mathbf{X}^* - \mathbf{X}_n)\right] = o(T)$, i.e. (53) leads to no regret (a.s.), as claimed.

Finally, for the mean regret bound (34), taking the expectation of the first line of (55) yields:

$$\mathbb{E}[\mathrm{Reg}(T)] \leq \frac{1}{\gamma_T} + \frac{\hat{V}_0^2}{2} \sum_{n=1}^{T} \gamma_n, \quad (58)$$

where we used the fact that $\mathrm{Reg}(T) \leq \mathbb{E}[\mathbf{V}_n \cdot (\mathbf{X}^* - \mathbf{X}_n)] = \mathbb{E}[\hat{\mathbf{V}}_n \cdot (\mathbf{X}^* - \mathbf{X}_n)]$ for the LHS, and the assumption that $\mathbb{E}[\|\hat{\mathbf{V}}_n\|^2] \leq \hat{V}_0^2$ for the RHS. ■

Proof of Proposition 2: By reasoning as in the proof of Proposition 1, we readily obtain:

$$\sum_{n=1}^{T} \mathrm{tr}[\hat{\mathbf{V}}_n \cdot (\mathbf{X}^* - \mathbf{X}_n)] \leq \frac{1}{2} \sum_{n=1}^{T} \gamma_n \|\hat{\mathbf{V}}_n\|^2, \quad (59)$$

so (37) follows by taking expectations on both sides as in the proof of Theorem 2. That $\mathbf{X}_n$ leads to no regret then follows by noting that (57) holds even for $\alpha = 1$. ■

REFERENCES

[1] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[2] Qualcomm, "The 1000x data challenge," White Paper, 2013.

[3] Huawei Technologies, "5G: A technology vision," White paper, 2013.

[4] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE Trans. Wireless Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

[5] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[6] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[7] P. Coucheney, B. Gaujal, and P. Mertikopoulos, "Distributed optimization in multi-user MIMO systems with imperfect and delayed information," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT'14)*, 2014, pp. 3097–3101.

[8] P. Mertikopoulos and A. L. Moustakas, "Learning in an uncertain world: MIMO covariance matrix optimization with imperfect feedback," *IEEE Trans. Signal Process.*, vol. 64, no. 1, pp. 5–18, Jan. 2016.

[9] D. Feng, C. Jiang, G. Lim, L. J. Cimini Jr., G. Feng, and G. Y. Li, "A survey of energy-efficient wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 167–178, Feb. 2013.

[10] G. Miao, N. Himayat, and G. Y. Li, "Energy-efficient link adaptation in frequency-selective channels," *IEEE Trans. Commun.*, vol. 58, no. 2, pp. 545–554, Feb. 2010.

[11] G. Miao, N. Himayat, G. Y. Li, and S. Talwar, "Distributed interference-aware energy-efficient power optimization," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1323–1333, Apr. 2011.

[12] C. Isheden, Z. Chong, E. Jorswieck, and G. Fettweis, "Framework for link-level energy efficiency optimization with informed transmitter," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2946–2957, Aug. 2012.

[13] X. Ge et al., "Energy efficiency optimization for MIMO-OFDM mobile multimedia communication systems with QoS constraints," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2127–2138, Jun. 2014.

[14] G. Bacci, E. V. Belmega, P. Mertikopoulos, and L. Sanguinetti, "Energy-aware competitive power allocation for heterogeneous networks under QoS constraints," *IEEE Trans. Wireless Commun.*, vol. 14, no. 9, pp. 4728–4742, Sep. 2015.

[15] Z. Wang, I. Stupia, and L. Vandendorpe, "Energy efficient precoder design for MIMO-OFDM with rate-dependent circuit power," in *Proc. IEEE Int. Conf. Commun.*, 2015, pp. 1897–1902.

[16] E. V. Belmega and S. Lasaulce, "Energy-efficient precoding for multiple-antenna terminals," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 329–340, Jan. 2011.

[17] Y. J. A. Zhang and K. B. Letaief, "An efficient resource-allocation scheme for spatial multiuser access in MIMO/OFDM systems," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 107–116, Jan. 2005.

[18] F. Meshkati, H. V. Poor, S. C. Schwartz, and N. B. Mandayam, "An energy-efficient approach to power control and receiver design in wireless data networks," *IEEE Trans. Commun.*, vol. 53, no. 11, pp. 1885–1894, Nov. 2005.

[19] F. Meshkati, M. Chiang, H. V. Poor, and S. C. Schwartz, "A game-theoretic approach to energy-efficient power control in multicarrier CDMA systems," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 6, pp. 1115–1129, Jun. 2006.

[20] E. V. Belmega, S. Lasaulce, and M. Debbah, "A survey on energy-efficient communications," in *Proc. 21st Annu. IEEE Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC'10)*, 2010, pp. 289–294.

[21] E. Eraslan, C.-Y. Wang, and B. Daneshrad, "Practical energy-aware link adaptation for MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 246–258, Jan. 2014.

[22] X. Huang, X. Ge, Y. Wang, F. Y. Li, and J. Zhang, "Energy-efficient binary power control with bit error rate constraint in MIMO-OFDM wireless communication systems," in *Proc. IEEE Veh. Technol. Conf. (VTC-Fall)*, 2012, pp. 1–5.

[23] J. Hannan, "Approximation to Bayes risk in repeated play," in *Contributions to the Theory of Games, Volume III*, M. Dresher, A. W. Tucker, and P. Wolfe, Eds. Princeton, NJ, USA: Princeton Univ. Press, 1957, vol. 39, pp. 97–139.

[24] S. Shalev-Shwartz, "Online learning and online convex optimization," *Found. Trends Mach. Learn.*, vol. 4, no. 2, pp. 107–194, 2011.
[25] N. Cesa-Bianchi and G. Lugosi, *Prediction, Learning, and Games*. Cambridge, U.K.: Cambridge Univ. Press, 2006.
[26] E. Hazan, "A survey: The convex optimization approach to regret minimization," in *Optimization for Machine Learning*, S. Sra, S. Nowozin, S. J. Wright, Eds. Cambridge, MA, USA: MIT Press, 2012, pp. 287–304.
[27] S. Maghsudi and S. Stanczak, "Joint channel selection and power control in infrastructureless wireless networks: A multi-player multi-armed bandit framework," *IEEE Trans. Veh. Technol.*, vol. 64, no. 10, pp. 4565–4578, Jul. 2014.
[28] J. Dams, M. Hoefer, and T. Kesselheim, "Convergence time of power-control dynamics," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 11, pp. 2231–2237, Dec. 2012.
[29] A. Blum and Y. Mansour, "From external to internal regret," *J. Mach. Learn. Res.*, vol. 8, pp. 1307–1324, Dec. 2007.
[30] I. Stiakogiannakis, P. Mertikopoulos, and C. Touati, "No regrets: Distributed power control under time-varying channels and QoS requirements," in *Proc. 51st Annu. Allerton Conf. Commun. Control Comput. (Allerton'14)*, 2014, pp. 213–220.
[31] P. Mertikopoulos and E. V. Belmega, "Transmit without regrets: Online optimization in MIMO–OFDM cognitive radio systems," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 11, pp. 1987–1999, Nov. 2014.
[32] I. Stiakogiannakis, P. Mertikopoulos, and C. Touati, "Adaptive power allocation and control in time-varying multi-carrier MIMO networks," 2015, http://arxiv.org/abs/1503.02155.
[33] S. M. Kakade, S. Shalev-Shwartz, and A. Tewari, "Regularization techniques for learning with matrices," *J. Mach. Learn. Res.*, vol. 13, pp. 1865–1890, 2012.
[34] K. Tsuda, G. Rätsch, and M. K. Warmuth, "Matrix exponentiated gradient updates for on-line Bregman projection," *J. Mach. Learn. Res.*, vol. 6, pp. 995–1018, 2005.
[35] M. Zinkevich, "Online convex programming and generalized infinitesimal gradient ascent," in *Proc. 20th Int. Conf. Mach. Learn. (ICML'03)*, 2003, pp. 928–936.
[36] E. V. Belmega and P. Mertikopoulos, "Energy-efficient power allocation in dynamic multi-carrier systems," in *Proc. IEEE Veh. Technol. Conf. (VTC-Spring)*, Glasgow, Scotland, May 2015, pp. 1–5.
[37] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun. Related Technol.*, vol. 10, no. 6, pp. 585–596, 1999.
[38] S. N. Diggavi and T. M. Cover, "The worst additive noise under a covariance constraint," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 3072–3081, Nov. 2001.
[39] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1089–1098, Aug. 2004.
[40] M. Hata, "Empirical formula for propagation loss in land mobile radio services," *IEEE Trans. Veh. Technol.*, vol. VT-29, no. 3, pp. 317–325, Aug. 1980.
[41] COST Action 231, "Digital mobile radio towards future generation systems," European Commission, Brussels, Belgium, Final report no. EUR 18957, 1999.
[42] J. Hofbauer and W. H. Sandholm, "Stable games and their dynamics," *J. Econ. Theory*, vol. 144, pp. 1710–1725, 2009.
[43] N. Cesa-Bianchi, P. Gaillard, G. Lugosi, and G. Stoltz, "Mirror descent meets fixed share (and feels no regret)," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, vol. 25, pp. 989–997.
[44] S. Hart and A. Mas-Colell, "A simple adaptive procedure leading to correlated equilibrium," *Econometrica*, vol. 68, no. 5, pp. 1127–1150, Sep. 2000.
[45] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1986–1992, Nov. 1997.
[46] S. Vishwanath, N. Jindal, and A. J. Goldsmith, "Duality, achievable rates and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.
[47] A. Charnes and W. W. Cooper, "Programming with linear fractional functionals," *Naval Res. Logist. Quart.*, vol. 9, pp. 181–196, 1962.
[48] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[49] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 145–152, Jan. 2004.
[50] G. Scutari, D. P. Palomar, and S. Barbarossa, "The MIMO iterative waterfilling algorithm," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1917–1935, May 2009.
[51] A. M. Davie and A. J. Stothers, "Improved bound for complexity of matrix multiplication," *Proc. Roy. Soc. Edinb. A*, vol. 143, no. 2, pp. 351–369, 2013.
[52] N. Maculan and G. G. de Paula Jr., "A linear-time median-finding algorithm for projecting a vector on the simplex of $\mathbb{R}^n$," *Oper. Res. Lett.*, vol. 8, pp. 219–222, 1989.
[53] J. Kwon and P. Mertikopoulos, "A continuous-time approach to online optimization," 2014, http://arxiv.org/abs/1401.6956.
[54] E. Hazan, A. Agarwal, and S. Kale, "Logarithmic regret algorithms for online convex optimization," *Mach. Learn.*, vol. 69, nos. 2–3, pp. 169–192, Dec. 2007.
[55] M. Stojanovic, J. G. Proakis, and J. A. Catipovic, "Analysis of the impact of channel estimation errors on the performance of a decision-feedback equalizer in fading multipath channels," *IEEE Trans. Commun.*, vol. 43, nos. 2–4, pp. 877–886, Feb./Apr. 1995.
[56] A. Pascual-Iserte, D. P. Palomar, A. I. Pérez-Neira, and M. Á. Lagunas, "A robust maximin approach for MIMO communications with partial channel state information based on convex optimization," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 346–360, Jan. 2006.
[57] L. Zhang, Y.-C. Liang, Y. Xin, and H. V. Poor, "Robust cognitive beamforming with partial channel state information," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4143–4153, Aug. 2009.
[58] J. Wang, G. Scutari, and D. P. Palomar, "Robust MIMO cognitive radio via game theory," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1183–1201, Mar. 2011.
[59] Y. Zhang, E. Dall'Anese, and G. B. Giannakis, "Distributed optimal beamformers for cognitive radios robust to channel uncertainties," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6495–6508, Dec. 2012.
[60] Y. Yang, G. Scutari, P. Song, and D. P. Palomar, "Robust MIMO cognitive radio under interference temperature constraints," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2465–2483, Nov. 2013.
[61] I. Stiakogiannakis, D. Zarbouti, G. Tsoulos, and D. Kaklamani, "Subcarrier allocation algorithms for multicellular OFDMA networks without channel state information," in *Proc. 3rd Int. Symp. Wireless Pervasive Comput. (ISWPC'08)*, 2008, pp. 73–77.
[62] 3GPP, "User equipment (UE) radio transmission and reception," White paper, Jun. 2014.
[63] L. Yang, T. Yafei, and Y. Chenyang, "Energy-efficient coordinated beamforming with individual data rate constraints," in *Proc. IEEE 24th Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, 2013, pp. 1040–1044.
[64] P. Hall and C. C. Heyde, *Martingale Limit Theory and Its Application*. New York, NY, USA: Academic, 1980.

Panayotis Mertikopoulos (M'11) received the "Ptychion" (diploma) degree in physics from the University of Athens, Athens, Greece, in 2003, the M.Sc. and M.Phil. degrees in mathematics from Brown University, Providence, RI, USA, in 2005 and 2006, respectively, and the Ph.D. degree in physics from the University of Athens, in 2010. From 2010 to 2011, he was a Postdoctoral Researcher with the Economics Department, École Polytechnique, Paris, France. Since 2011, he has been a CNRS Researcher with the Laboratoire d'Informatique de Grenoble, Grenoble, France. His research interests include dynamical systems, optimization, game theory, and their applications to wireless networks. He has been an Embeirikeion Foundation Fellow, since 2003. He was a track Co-Chair of IFORS 2014, TPC Co-Chair of WiOpt 2014, and General Co-Chair of AlgoGT 2013. He was the recipient of the Best Paper Award in NetGCoop'12.

E. Veronica Belmega (S'08–M'10) received the M.Sc. (engineer diploma) degree from the University Politehnica of Bucharest, Bucharest, Romania, in 2007, and the M.Sc. and Ph.D. degrees both from the Université Paris-Sud 11, Orsay, France, in 2007 and 2010, respectively. From 2010 to 2011, she was a Postdoctoral Researcher in a joint project between the Alcatel-Lucent Chair on Flexible Radio in Supélec and Princeton University, Princeton, NJ, USA. She is currently with ETIS/ENSEA—Université de Cergy-Pontoise—CNRS, Cergy-Pontoise, France, and with Inria, France. She was one of the ten recipients of the 2009 L'Oréal—UNESCO—French Academy of Science Fellowship: "For young women doctoral candidates in science."

Online Power Optimization in Feedback-Limited, Dynamic and Unpredictable IoT Networks

Alexandre Marcastel, *Student Member, IEEE*, E. Veronica Belmega, *Member, IEEE*, Panayotis Mertikopoulos, *Member, IEEE*, and Inbar Fijalkow, *Senior Member, IEEE*

***Abstract*—One of the key challenges in Internet of Things (IoT) networks is to connect many different types of autonomous devices while reducing their individual power consumption. This problem is exacerbated by two main factors: *a*) the fact that these devices operate in and give rise to a highly dynamic and unpredictable environment where existing solutions (e.g., water-filling algorithms) are no longer relevant; and *b*) the lack of sufficient information at the device end. To address these issues, we propose a regret-based formulation that accounts for *arbitrary network dynamics*: this allows us to derive an online power control scheme which is provably capable of adapting to such changes, while relying solely on *strictly causal feedback*. In so doing, we identify an important tradeoff between the amount of feedback available at the transmitter side and the resulting system performance: if the device has access to unbiased gradient observations, the algorithm's regret after T stages is $\mathcal{O}(T^{-1/2})$ (up to logarithmic factors); on the other hand, if the device only has access to scalar, utility-based information, this decay rate drops to $\mathcal{O}(T^{-1/4})$. The above is validated by an extensive suite of numerical simulations in realistic channel conditions, which clearly exhibit the gains of the proposed online approach over traditional water-filling methods.**

***Index Terms*—IoT networks, online exponential learning, imperfect and scarce feedback**

I. Introduction

THE emerging Internet of things (IoT) paradigm is projected to connect billions of wireless "things" (wireless sensors, wearables, biochip transponders, etc.) in a vast network with drastically different requirements between components (e.g. in terms of throughput and power characteristics) [3]. Following Moore's prediction on silicon integration, the wireless surroundings of IoT networks are expected to exhibit massive device densities with high interference levels. An orthogonal spectrum allocation is therefore energetically inefficient, as an unrealistic number of bands or subcarriers would be required to accommodate all devices. The usage of new access protocols such as non-orthogonal multiple access (NOMA) [4] is considered instead, in which interference mitigation becomes critical. For this reason, and also given that the autonomous wireless devices have stringent battery limitations, optimizing the power consumption emerges as one of the key ingredients for achieving a "speed of thought" user experience at the application level [5].

A major challenge that arises here is that IoT networks are characterized by an unprecedented degree of temporal variability – due itself to the unique mobility attributes of modern wearable devices, intermittent user activity, application diversity etc. As such, IoT networks cannot be treated as static (or stationary) systems, implying in turn that conventional optimization techniques that *target a fixed state,* for instance via water-filling type of algorithms, are no longer relevant. The main limitation of classical approaches is their lack of robustness to strictly causal – *no look-ahead* – channel state information, which is inevitable in dynamic, unpredictable environments. Therefore, power optimization in dynamic IoT networks calls for a different toolbox that is provably capable of adapting to unpredictable changes in the network.

Motivated by its prolific success in the fields of machine learning and artificial intelligence [6, 7], we propose in this paper a *regret-based* formulation of power optimization which allows us to consider arbitrary variations in the network. The core component of this approach is that, instead of targeting a specific network state, it aims to derive an online power allocation policy whose performance over time is as close as possible to that of the best fixed policy in hindsight (even though computing the latter requires non-causal knowledge of the system parameters and their evolution ahead of time). Owing to this straightforward and flexible definition, regret minimization has become the leading paradigm for online decision making in uncertain, dynamic environments, ranging from online ad auctions [8, 9] and recommender systems [10] to throughput and energy efficiency optimization problems in wireless communications [11, 12].

A critical performance limitation in the above is the fact that wireless devices in IoT networks typically receive limited and/or corrupted feedback from their environment [13]. To name but an example, channel state information (CSI) is usually acquired by the access point (AP) using

A. Marcastel, E. V. Belmega and I. Fijalkow are with ETIS, Université Paris Seine, Université Cergy-Pontoise, ENSEA, CNRS, Cergy-Pontoise, France. P. Mertikopoulos is with Univ. Grenoble Alpes, CNRS, Inria, LIG, Grenoble, France.

This research was supported in part by the Orange Lab Research Chair on IoT within the University of Cergy-Pontoise, by the French National Research Agency (ANR) project ORACLESS (ANR–16–CE33–0004–01), by the ELIOT ANR-18-CE40-0030 and FAPESP 2018/12579-7 project, and by ENSEA, Cergy-Pontoise, France. Part of this work was presented in VTC2016-Fall [1] and GLOBECOM 2016 [2].

pilot transmissions that are subsequently fed back to each device. Since IoT networks bring together massive numbers of devices, the signaling overhead increases to the point where it cannot be distributed over multiple frequency bands in an efficient manner (due to spectrum scarcity) [14, 15]. Therefore, to reduce the impact of this overhead, the amount of information fed back to wireless devices must be reduced as much as possible, and the resulting estimation errors must be likewise taken into account. The same kind of problem has been underlined in cooperative multi-user networks [16], in which the global network optimum objective leads to massive signaling; and in massive multiple-input and multiple-output (MIMO) systems [17–20], in which the increase in the number of antennas leads to a prohibitive amount of required CSI. Instead, in IoT networks, it is not the number of antennas but the large number of connected devices that create this bottleneck.

A. Summary of contributions and paper outline

In the field of online learning, the challenges that result from incomplete and/or imperfect feedback have been studied extensively in the context of the so-called multi-armed bandit problems [6]. These problems are inherently discrete in nature, so the lessons learned from this literature do not apply to the power allocation framework studied here (a continuous, multi-dimensional problem in itself). Nevertheless, by leveraging ideas originating in the well-known exponential weights algorithm for multi-armed bandits [6], we derive an online power allocation policty based on exponentiated gradient descent (EGD), and which comprises two basic steps: *a*) tracking the gradient of the users' power minimization objective in a dual, unconstrained space; and *b*) using a judiciously designed exponential function to map the output of this step to a feasible power allocation profile and keep going. Thanks to this two-step, primal-dual approach, we are then able to derive concrete regret minimization guarantees for the online power minimization problem, irrespective of the network's dynamics.

To establish a benchmark, we begin with the *full information* or the first-order feedback case, where each wireless device is assumed to have perfect feedback on the gradient of its individual power minimization objective. In this case, the proposed power allocation policy is shown in Section III to enjoy a $\mathcal{O}(T^{-1/2})$ regret guarantee, meaning that the algorithm's performance over a horizon of T transmission cycles is no more than $\mathcal{O}(T^{-1/2})$ away from the best fixed policy in hindsight. Importantly, unless rigid statistical hypotheses are made for the underlying IoT network (such as assuming that it evolves following a stationary ergodic process), this guarantee cannot be improved; however, we show in Section IV that it can still be attained even if the feedback received by each device is imperfect and/or otherwise corrupted by non-systematic measurement errors and observational noise.

In addition to providing a comparison baseline, the full information case also allows us to compare the performance of the proposed algorithm to that of classical water-filling algorithms [21–23] and highlight the difficulties encountered by the latter when the network evolves dynamically over time and only a strictly causal (with no look-ahead) feedback information is available at the transmitter.

On the other hand, if the only information received by each device is the observed value of their power minimization objective (the so-called *zeroth-order feedback* setting), these bounds change significantly. Lacking any sort of vector-valued, gradient-based feedback, we rely on simultaneous stochastic approximation techniques [6, 7], to build an estimator for the gradient: importantly, this estimator is potentially biased, but its bias can be controlled by tuning a certain sampling parameter. By jointly optimizing the value of this parameter and that of the original algorithm's step-size, we then show that the proposed policy still leads to no regret, but now at a slower rate of $\mathcal{O}(T^{-1/4})$.

In Section VI, we validate our theoretical analysis via numerical experiments and highlight highly dynamic networks with realistic, unpredictable channel conditions. Classical water-filling algorithms are very sensitive to unpredictable changes in the network and are outperformed by our proposed online algorithms in terms of power consumption and achieved rate. Concerning the impact of available feedback, our numerical results also illustrate a compromise between the amount and/or quality of the feedback information and the algorithms' performance (measured here in terms of the time needed to attain a no-regret state). The zeroth-order feedback case requires only the knowledge of a scalar at each iteration (the value of the objective function) as opposed to a vector (the gradient), but the average time required to reach a no-regret state is higher.

B. Related works

Regarding resource allocation in static IoT environments, several problems have been studied [24–26]. In [24] the authors study the resource allocation for machine to machine (M2M) communications using cooperative game theoretic tools in which the machines want to maximize their own rate. In [25], the authors study the problem of clustering and power allocation for both uplink and downlink in NOMA systems. Similar to [24], each device aims at maximizing its own rate. To solve this problem, the authors used classical optimization tools. In [26], the problem of power control for mutual interference avoidance is studied by using also classical optimization tools. In all these works, the network is assumed to remain static over time and the devices are required to have perfect feedback information. Here, we relax both assumptions by taking into account the inherent dynamics of an IoT network and the impact of feedback imperfections and scarcity.

In (non-IoT) wireless networks, there exists a wide resource allocation literature essentially concerned with either static [21–23, 27, 28] or stochastic [16, 20, 29–35] optimization problems, to cite but a few. In these works, the

underlying network is assumed to remain static or to evolve following a stationary random process. Their main aim is to derive efficient algorithms, based on classical optimization, stochastic optimization, or on machine learning tools, that converge to an optimal fixed or steady state. These works are inherently different from the present paper, in which we squarely focus on arbitrarily dynamic networks (the network can even evolve in a non-stationary way). In such unpredictable networks, there is no fixed *solution state* to converge to, so the very notion of *convergence* as a performance metric needs to be rethought from the ground up.

Adaptive allocation policies based on online optimization tools have been recently proposed but in quite different settings and problems [11, 12, 36–38]. In [36], the authors proposed a multi-armed bandit formulation of the channel selection problem and derived an online channel selection algorithm using upper confidence bound techniques; a similar approach has also been used in the context of beam-alignment for millimeterWave communications [37]. In [38], an adversarial multi-armed bandit formulation is proposed to tackle an access point association problem in hybrid indoor LiFi-WiFi communication systems exploiting the exponential weights algorithm. For IoT networks, the recent work [39] acknowledges the high potential of the online learning framework and then focuses on multi-armed bandits for mobile computation offloading problems at the edge layer. However, in our setting, the agents' decisions are not taken within a stochastic environment (so upper confidence bounds are not applicable) and all variables are continuous as opposed to discrete (so multi-armed bandits are not suitable).

Regarding physical-layer resource allocation problems, the authors of [11, 12] studied dynamic MIMO systems from the point of view of online throughput and energy efficiency maximization. By contrast, our focus here is the power minimization problem in IoT networks, which is inherently different. Specifically, in the online throughput maximization problem in [11], the opportunistic devices have to always transmit at full available power, which is not power-efficient and the proposed learning algorithm does not apply to the problem at hand. The energy efficiency (defined as the ratio between the achieved rate and the overall power consumption) maximization problem in [12] is non-convex and is cast into a convex problem by performing a suitable variable change, which results into a specific exponential learning algorithm that also does not apply here. However, the learning algorithms in these works rely on the availability of gradient information which amounts to a (typically large) matrix worth of feedback; by contrast, the algorithm provided in this paper only requires a single readily available scalar as feedback at the device end.

To the best of our knowledge, our paper is the first in the IoT literature to take into account the network's inherent dynamics and its unpredictable temporal variability when designing power-efficient allocation policies. Furthermore, it is among the first works in the resource allocation literature in multi-user wireless networks, proposing an adaptive algorithm relying only on a single scalar feedback information (the sole past experienced objective value).

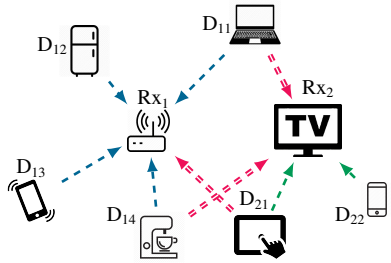


Figure 1: System composed of six transmit devices (D_{11}, D_{12}, etc.) and two receivers (Rx_1, Rx_2). The blue and green arrows represent the direct links while the red (double-lined) ones are interfering links.

II. Model and Problem Formulation

We consider a system composed of M transmitters and N receivers communicating over S orthogonal subcarriers or sub-bands as illustrated in Fig. 1: each device transmits to only one intended receiver, but a given receiver may decode several incoming signals.

Since we aim at devising a distributed policy that needs no central controller, we can focus on one particular transmitting-receiving pair. The received signal for the (arbitrarily chosen) focal device becomes:

$$y^s(t) = h^s(t)x^s(t) + \sum\nolimits_j h_j^s(t)x_j^s(t) + z^s(t), \qquad (1)$$

where $s \in \{1, \ldots, S\}$ is the subcarrier index; $x^s(t)$ is the transmitted signal; $h^s(t)$ is the channel gain between the focal transmitter and its intended receiver; x_j^s is the transmitted signal of device j; $h_j^s(t)$ is the interfering channel gain between transmitter j and the focal receiver; and $z^s(t)$ is the received noise of the focal device.

We also define the effective channel gain vector $\mathbf{w}(t) = (w^s(t))$, where $w^s(t)$ represents the effective gain in subcarrier s and is given by

$$w^s(t) = \frac{g^s(t)}{\sigma^2 + \sum_j g_j^s(t)p_j^s(t)}, \quad \forall s, \qquad (2)$$

where σ^2 is the variance of the noise $z^s(t)$, $p_j^s(t)$ is the transmitted power by the user j in subcarrier s, $g_j^s(t) = |h_j^s(t)|^2$ and $g^s(t) = |h^s(t)|^2$.

The above expression implies that the receiver employs single-user decoding (SUD), meaning that, when decoding a transmitted signal, the other incoming signals are treated as noise. This consideration is relevant in distributed and energy-limited networks, such as IoT networks, in which the receivers may not be able to decode the interfering signals (e.g., may not know the codebooks of the interferers). Also, the receivers may not afford to sequentially process and decode their incoming signals (via successive interference

cancellation) and the transmitting devices may not be coordinated (and may not know their decoding order).

The aim of IoT networks is to interconnect many different types of devices in a distributed or self-optimizing way. Since most of them are likely to be small devices like sensors, phones, or isolated devices that operate solely on limited batteries, reducing the power consumption is a key challenge in IoT networks [40, 41]. In this work, the main objective is to minimize the power consumption of the focal device taking into account its quality of service (QoS) requirements. These requirements are performance targets, which depend on the specific application, for example: a) minimum rate per device (i.e., the rate of the focal device has to be higher than a given threshold R_{min}); b) minimum SINR per device; c) minimum network sum-rate. Such QoS requirements differ from physical hard constraints (e.g., the transmit power positivity constraints) in that they cannot be ideally guaranteed - always at 100% - in practical communication systems and some outage has to be generally tolerated.

In view of the above, the trade-off between power minimization and QoS requirements will be modeled via the loss function:

$$L_t(\mathbf{p}) = \sum_{s=1}^{S} p^s \quad + \lambda \ [R_{min} - R_t(\mathbf{p})]^+ \tag{3}$$

where $\mathbf{p} = (p^1, \ldots, p^S)$ represents the power allocation vector of the focal device with components $p^s, \forall s$ representing the power allocated to the s-th subcarrier. The first term in the objective is the overall power consumption and the second term is a soft-constraint (or penalty) term, which is activated whenever the minimum target rate R_{min} is not achieved. Finally, $R_t(\mathbf{p})$ denotes the well-known Shannon rate:

$$R_t(\mathbf{p}) = \sum_{s=1}^{S} \log(1 + w^s(t)p^s) \tag{4}$$

and $[x]^+ \triangleq \max\{x, 0\}$, meaning that no penalty is applied when the achieved rate is greater than the threshold $R_t(\mathbf{p}) \geq R_{min}$. Although we choose a linear penalty function for its relevance to communications [28, 42, 43] (and to simplify the presentation), our results carry over the more general class of concave functions, e.g., logarithmic penalties [44]. The parameter λ can also be interpreted as the unit-cost for each bps/Hz under the QoS target R_{min} and, as we will see in Section VI, it also represents a sensitivity parameter that has to be carefully tuned to adjust the flexibility regarding the minimum rate constraint violations or outages. Indeed, higher values of λ lead to less QoS outages, but at the cost of incurring a higher power consumption.

To sum up, the online optimization problem under study can be stated as:

$$\begin{array}{ll} \text{minimize} & L_t(\mathbf{p}(t)) \\ \text{over} & \mathbf{p}(t) = (p^1(t), \ldots, p^S(t)) \\ \text{subject to} & p^j(t) \geq 0, \ \ \forall j \in \{1, \ldots, S\} \\ & \sum_{s=1}^{S} p^s(t) \leq P_{\max} \ . \end{array} \tag{5}$$

The minimization variable is the power allocation vector of the focal device across the available frequency subcarriers, $\mathbf{p}(t)$, and both constraints are physical ones. The first constraint guarantees that the transmit power of the focal device, in each subcarrier j, is always positive. The second constraint comes from the power supply limitation and implies that the total power of the focal device that is spread over the subcarriers is bounded from above by the maximum transmit power of the device.

Concerning the above objective function $L_t(\mathbf{p})$, notice that it *may vary in a non-stationary and unpredictable way* such that the focal device cannot determine *a priori* (before the transmission takes place) its instantaneous or dynamic optimal power allocation $\mathbf{p}^*(t)$ that minimizes this objective at each time t. Nevertheless, we assume that the device receives some feedback after each transmission, such as the past experienced objective value or its past gradient. The idea in online optimization is to exploit this strictly causal feedback information to build a dynamic and adaptive power allocation policy $\mathbf{p}(t)$ that *minimizes as much as possible* the time-varying objective function $L_t(\mathbf{p}(t))$[1].

The major novelty in the above formulation relative to more classical power allocation problems lies in its dynamic nature and the fact that we make *no assumptions* on the network dynamics. Notice that, the objective function in (3) depends on the network dynamics via the second penalty term. Indeed, the achieved rate $R_t(\mathbf{p}(t))$ depends on the varying wireless channels and also on the power allocation policies of the other devices via the effective channel gains in (2). Classical approaches leading to water-filling type of algorithms [21–23] rely both on static (or stationary) channel models and on strong assumptions on the information available at the transmitter before the transmission takes place (e.g., perfect channel state information in the form of the SINR in each subcarrier). In highly dynamic and distributed IoT networks, such assumptions are too stringent and no longer hold.

On that account, the aim of this work is twofold: to explicitly take into account the device mobility, their network connectivity patterns and behaviour, which may be completely arbitrary and unpredictable; and to greatly reduce the information required at the transmitter.

In order to evaluate the performance of a given online

[1]Going back to our model of the QoS requirement, another motivation for including it into the objective function in (3), as opposed to imposing a hard constraint, is that the latter would result in an arbitrarily time-varying and unpredictable feasible set at the decision instant. This issue is highly non-trivial and open in online optimization, which would require going well beyond the standard regret minimization framework and, hence, falls out of the scope of this work.

policy $\mathbf{p}(t)$, the most commonly used notion is that of the regret [6, 7, 11, 12], which compares its performance in terms of loss with a benchmark policy. Now, comparing any policy $\mathbf{p}(t)$, built using outdated feedback information, with the instantaneous or dynamic optimal solution $\mathbf{p}^*(t)$ is obviously too ambitious. Instead, the notion of regret compares the policy $\mathbf{p}(t)$ to a less ambitious benchmark: the fixed strategy that minimizes the overall objective over a given transmission horizon T:

$$Reg(T) \triangleq \sum_{t=1}^{T} L_t(\mathbf{p}(t)) - \min_{\mathbf{q}\in\mathcal{P}} \sum_{t=1}^{T} L_t(\mathbf{q}), \tag{6}$$

where $\mathcal{P} \triangleq \left\{\mathbf{p} \in \mathbb{R}^S \;\middle|\; p^s \geq 0, \forall s, \sum_{s=1}^{S} p^s \leq P_{max}\right\}$ denotes the feasible set. Otherwise stated, the regret measures the performance gap between a power allocation policy $\mathbf{p}(t)$ and the best mean optimal solution over a fixed horizon T. If the regret is negative, then the dynamic policy $\mathbf{p}(t)$ outperforms the best mean optimal solution overall. To quantify this, the policy $\mathbf{p}(t)$ is said to lead to no-regret if

$$\limsup_{T\to\infty} \frac{1}{T} Reg(T) \leq 0. \tag{7}$$

A no-regret policy $\mathbf{p}(t)$ is asymptotically optimal and performs at least as good as the best fixed strategy on average (when T grows large).

Notice that although the best mean optimal solution is less ambitious than $\mathbf{p}^*(t)$ (minimizing the objective function at each t) its computation requires the same non-causal knowledge of the system parameters and the evolution of the objective throughout the time horizon T before the transmission takes place, or in hindsight. Therefore, the design of dynamic policies that reach no-regret while relying on strictly causal and local information is a remarkable and desirable goal. Moreover, in the particular case of a static network composed of a single transmit device, the online optimization problem in (5) reduces to a classic convex optimization problem. A no-regret online policy $\mathbf{p}(t)$ in this case implies the convergence of the average policy: $\overline{\mathbf{p}}(t) \triangleq \frac{1}{T}\sum_{t=1}^{T} \mathbf{p}(t)$ to the solution set of the relevant optimization problem [45]. In conclusion, given the IoT network dynamics and unpredictability, our focus in the following is precisely to develop no-regret online policies for the online optimization problem defined in (5).

III. First-Order feedback

In the resource allocation problem under study, the focal device has to choose in which of the available subcarriers to transmit, how much of the available power to consume and how to split this amount over the chosen subcarriers, all this based on the strictly causal feedback information. This is reminiscent of the well-known multi-armed bandit problem in sequential online learning [7]: there, at each instant, the player chooses an action (or an arm) out of several possibilities and receives a reward as a result. Outside the so-called "stochastic" case (where each arm's payoff is determined by a fixed probability distribution), the most widely used algorithmic scheme is the exponential (or multiplicative) weights algorithm [7], where payoffs are aggregated over time and the optimizer selects an arm with a probability proportional to the exponential of these scores. In what follows, we derive the necessary machinery to extend this idea to the continuous optimization problem at hand and derive an exponentiated gradient descent algorithm for power minimization in this context.

OXL algoritm: Online Exponential Learning Algorithm

Initialization: $\mathbf{y}(0) \leftarrow 0$; $t \leftarrow 0$.

Repeat

- ▹ Pre-transmission phase: update transmit powers
- $\mathbf{p}(t) \leftarrow \mathbf{Q}(\mathbf{y}(t))$ defined in (OXL)
- ▹ Transmit at $\mathbf{p}(t)$
- ▹ Post-transmission phase: receive gradient feedback $\mathbf{v}(t)$
- Update scores $\mathbf{y}(t+1) \leftarrow \mathbf{y}(t) - \mu(t)\,\mathbf{v}(t)$
- $t \leftarrow t+1$

until transmission ends

In our setup, we begin by assuming that each device has access to some feedback mechanism that provides the first-order gradient information $\mathbf{v}(t) = \nabla L_t(\mathbf{p}(t))$ at the end of each transmission. Our proposed algorithm can be summarized in two steps. First, the device tracks the past gradient of its objective without taking account the power constraints. Second, the device maps the first step into the feasible set $\mathcal{P}$ using a well chosen exponential map as follows:

$$\begin{aligned} &\mathbf{y}(t) = \mathbf{y}(t-1) - \mu\,\mathbf{v}(t), \\ &p^s(t) = Q^s(\mathbf{y}(t)) \triangleq P_{\max} \frac{\exp(y^s(t))}{1 + \sum_{i=1}^{S} \exp(y^i(t))}, \end{aligned} \tag{OXL}$$

where μ is the step-size parameter. We denote by $\mathbf{Q}(\mathbf{y}(t)) = (Q^1(t), \ldots, Q^S(t))$ the exponential vector field that maps the updated score $\mathbf{y}(t)$ into the feasible set.

Essentially, the online exponential learning algorithm detailed above, tracks the cumulative negative gradient of the convex loss function and then maps the result to the feasible set. The exponential mapping step could be replaced by an Euclidean projection and the resulting algorithm would be an online gradient descent [46] algorithm. We chose the exponential mapping because of its *reduced complexity* relative to a projected gradient descent algorithm that would require an additional (possibly costly) projection step. Indeed, from (OXL) it is easy to see that the updates are easy to compute and that they meet the contraints. More precisely, the complexity of each iteration t is linear in the problem dimensionality S, the number of subcarriers over which the focal device transmits. Hence, given that S is not expected to grow large for a specific IoT device (transmitting on a small subset of the total number of subcarriers available to the entire IoT network), the OXL algorithm is particularly appealing for distributed, device-centric IoT networks.

We will now study the evolution of the regret of the dynamic power allocation policy (OXL) to show that it

holds the no-regret property. To that end, let V denote an upper bound on the gradient feedback $\mathbf{v}(t)$ in the sense that $\|\mathbf{v}(t)\|^2 \leq V$. We then have the following result (for a proof, see Appendix A):

Theorem 1. *If the OXL algorithm is run with a constant step-size μ then, it enjoys the regret bound:*

$$Reg(T) \leq \frac{P_{\max} \log(1+S)}{\mu} + \frac{\mu P_{\max} TV}{2}. \tag{8}$$

Tuning the step-size μ: The step-size μ plays an important role in the *exploration vs. exploitation* tradeoff and, hence, in the ability of OXL algorithm to reach the no regret state, as we will see in the following. Intuitively, small values of μ imply that the subcarriers are almost equally explored and the gradient information is not exploited enough. High values of μ imply that only the best performing carriers w.r.t. past gradients are exploited and highly potential carriers, which have not performed well in the past, are rooted out too soon.

Notice that the above upper bound grows linearly with T, which may lead to a non-zero average regret. Nevertheless, this bound is a convex function of the step-size μ and, hence, can be minimized w.r.t. μ by setting the first-order derivative to zero. The resulting optimal step-size is

$$\mu^* = \sqrt{2\log(1+S)/(TV)}, \tag{9}$$

which then yields the sub-linear optimal bound

$$Reg(T) \leq P_{\max} \sqrt{2TV\log(1+S)}. \tag{10}$$

Therefore, by carefully choosing the step-size μ, OXL algorithm leads to no regret:
$\limsup_{T\to\infty} \frac{1}{T} Reg(T) = 0$.

Corollary 1. *If the OXL algorithm is run for a known horizon T using the optimal step-size μ^* in* (9)*, then it leads to no regret and the average regret $Reg(T)/T$ decays as $\mathcal{O}(T^{-1/2})$.*

The resulting regret bound in (10) depends on the system parameters: the total power $P_{\max}$, the number of subcarriers S, an upper bound on the gradient norm V, but also on the transmission horizon T, which the device does not necessarily know in advance. To avoid this limitation, we use the doubling trick [47]: the algorithm is run repeatedly starting with a unit-size window (number of iterations) and then doubling the window size at each new run until transmission ends. Hence, each widow size is known and the device can compute the corresponding optimal step μ^* (by replacing T with the window size in (9)). The bound in Corollary 1 applies in each window and the following result is proven in Appendix B.

Proposition 1. *If the OXL algorithm is run when the transmission time T is unknown by using the doubling trick with an optimal step-size for each window until the transmission ends (as in Figure 2), the regret enjoys the following bound:*

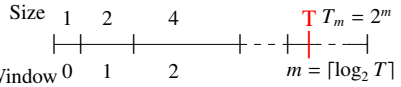


Figure 2: Illustration of the windows in the doubling trick.

$$Reg(T) \leq \frac{2}{\sqrt{2}-1} P_{\max} \sqrt{2TV\log(1+S)}. \tag{11}$$

Hence, OXL algorithm leads to no-regret and the average regret $Reg(T)/T$ decays as $\mathcal{O}(T^{-1/2})$.

We observe that not knowing the horizon T in advance results only in a small loss in the regret bound (in the multiplying constant).

IV. Imperfect gradient feedback

In this section, we relax the assumption of perfect gradient feedback and we consider that the focal device has access only to an imperfect gradient estimate, denoted by $\tilde{\mathbf{v}}(t)$, which meets the following conditions

$$\begin{aligned} \mathbb{E}[\tilde{\mathbf{v}}(t)] &= \nabla L_t(\mathbf{p}(t)), \\ \mathbb{E}[\|\tilde{\mathbf{v}}(t)\|^2] &\leq \tilde{V}, \end{aligned} \tag{12}$$

where the expectation is taken over the randomness of the estimator. These conditions are not very restrictive as they require the absence of systematic errors and a bounded variance, as such, they are satisfied by all common error distributions (Gaussian, log-normal, etc) [12]. For example, the common error model: $\tilde{\mathbf{v}}(t) = \nabla L_t(\mathbf{p}(t)) + \mathbf{z}$, where $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \sigma_z^2 \mathbf{I})$ [48] satisfies the above conditions.

Under these assumptions, the transmit powers in OXL algorithm are updated in function of $\tilde{\mathbf{v}}$ instead of the actual gradient $\mathbf{v}$ (via the internal score $\mathbf{y}(t)$). Thus, the online policy $\mathbf{p}(t)$ depends on the randomness of the estimator, which implies that the regret in (6) will also depend on this randomness. To take this into account, we study the *average regret* $\mathbb{E}[Reg(T)]$, where the expectation is taken over the randomness of the estimator. The no-regret property can be easily extended to the average regret as follows: a power allocation policy $\mathbf{p}(t)$ leads to no regret (on average) if

$$\limsup_{T\to\infty} \frac{1}{T} \mathbb{E}[Reg(T)] \leq 0. \tag{13}$$

A different possibility would be to study the probability that the regret in (6) falls bellow zero, but we leave this as a non-trivial open issue for future investigation.

Our first result, proved in Appendix C, concerns the case in which the transmission horizon T is known.

Theorem 2. *If the OXL algorithm is run for T iterations with a constant step-size μ and an imperfect gradient estimation defined in* (12)*, the average regret is bounded by:*

$$\mathbb{E}[Reg(T)] \leq \frac{P_{\max} \log(1+S)}{\mu} + \frac{\mu}{2} P_{\max} T \tilde{V}. \tag{14}$$

We observe that the upper bound above is somewhat similar to the one in Theorem 1 and can also be minimized with respect to the step-size μ for the same reasons. The optimal step-size is

$$\mu^* = \sqrt{2\log(1+S)/(T\tilde{V})}, \tag{15}$$

which provides the optimal bound

$$\mathbb{E}[Reg(T)] \leq P_{\max}\sqrt{2T\tilde{V}\log(1+S)}. \tag{16}$$

Therefore, using this optimal step-size leads to no regret even if the device only has access to imperfect gradient observations.

Corollary 2. *If the OXL algorithm is run for a known transmission time T with an imperfect gradient feedback and the optimal step-size μ^* in* (15)*, then the no-regret property holds and the average regret $\mathbb{E}[Reg(T)]/T$ decreases in $\mathcal{O}(T^{-1/2})$.*

The above result implies that *unbiased errors in the gradient estimation do not impact* greatly the evolution of the regret in expectation. This result should be contrasted to the corresponding one in the perfect gradient case. First, in the perfect gradient case, there is no randomness and the regret results are deterministic. Here, because of the random errors in the estimated gradient the above results hold only in expectation. Second, the obtained upper bounds depend on $\tilde{V}$, which is an upper bound on the second order statistics of the estimation over the entire time horizon (as opposed to the Lipschitz constant V). This means that the variance of the errors negatively impact the expected regret; higher variance errors result in higher expected regret.

Finally, if the device does not know in advance the transmission horizon T, the doubling trick described in Sec. III requires the knowledge of $\tilde{V}$ (to compute the optimal step-size). This may not be realistic in an unpredictable, time-varying and possibly non-stationary environment. To avoid this additional requirement, we take a variable step-size approach as in [6, 7, 49] and focus on the schedule

$$\mu(t) = \alpha/\sqrt{t}, \tag{17}$$

with parameter $\alpha > 0$. Using this variable step-size, we obtain the following result (for a proof, see Appendix D):

Theorem 3. *If the OXL algorithm is run with imperfect gradient feedback for an unknown horizon T and using the variable step-size $\mu(t) = \alpha t^{-1/2}$, then the average regret is bounded by*

$$\frac{\mathbb{E}[Reg(T)]}{T} \leq \frac{P_{\max}\log(1+S)}{\alpha\sqrt{T}} + \frac{P_{\max}\tilde{V}\alpha(1+\log T)}{2\sqrt{T}}. \tag{18}$$

Consequently, the device's average regret $\mathbb{E}[Reg(T)]/T$ vanishes as $\mathcal{O}(\log(T)T^{-1/2})$, i.e. OXL algorithm leads to no regret.

We remark the loss in the decay rate of the regret resulting from the lack of knowledge of $\tilde{V}$. This means that, with scarcer available knowledge, the device will reach the no regret state at a slower rate. Nevertheless, this loss is only logarithmic and even without the knowledge of T and relying on an imperfect and unbiased gradient feedback, the OXL algorithm is able to reach no regret.

V. Zeroth-order Feedback

In this section, our objective is to reduce even further the amount of required information to be fed back to the transmitting device. Instead of receiving a vector as feedback - the gradient or its unbiased estimation - the devices are now assumed to know only the value of the experienced objective function. This means that only a single scalar worth of information is needed at the transmitting device – a major advantage in feedback-limited and dynamic networks, where the acquisition of non-causal and complete channel state information (not to mention other network parameters) is a tall order. To the best of our knowledge, the proposed algorithm is the first adaptive power allocation algorithm for multiple-carrier, multiple-user networks requiring scalar feedback. Classic resource allocation algorithms such as water-filling policies require at least one quality indicator per subcarrier (e.g., the SINR value in each subcarrier), and, hence a (possibly large) vector worth of feedback.

To develop an online policy $\mathbf{p}(t)$ that leads to no regret, we modify the exponential mapping step in Sec. III and propose a novel learning algorithm that only requires zeroth-order feedback. To this aim, the first obstacle is to estimate the gradient of the objective based only on its value – in other words to do "gradient descent without a gradient" [50]. The main idea that we exploit here is the simultaneous stochastic approximation technique, which randomly samples the objective function in a neighbourhood of the power policy $\mathbf{p}(t)$ to obtain a *(potentially biased) estimate* of the gradient at this point [6, 7, 50].

For simplicity, we illustrate this technique on a particular directional derivative of $L_t(\mathbf{p})$ along the unit vector $\mathbf{x}$ (recall that the gradient is a collection of directional derivatives), denoted by $\nabla_{\mathbf{x}}L_t(\mathbf{p})$:

$$\nabla_{\mathbf{x}}L_t(\mathbf{p}) = \lim_{\delta\to 0}\frac{L_t(\mathbf{p}+\delta\mathbf{x}) - L_t(\mathbf{p}-\delta\mathbf{x})}{2\delta}, \tag{19}$$

which we want to estimate based on the single function value $L_t(\mathbf{p})$. To do so, we randomly sample the objective function around the point $\mathbf{p}$ in the direction $\mathbf{x}$ by drawing a Bernoulli distributed random variable $u \in \{-1, +1\}$ with equal probability. We can compute the expectation of these samples w.r.t. the randomness of u:

$$\mathbb{E}\left[L_t(\mathbf{p}+\delta u\mathbf{x})u\right] = \frac{L_t(\mathbf{p}+\delta\mathbf{x}) - L_t(\mathbf{p}-\delta\mathbf{x})}{2}. \tag{20}$$

From (19) and (20), we observe that

$$\mathbb{E}\left[\frac{L_t(\mathbf{p}+\delta u\mathbf{x})u}{\delta}\right] \approx \nabla_{\mathbf{x}}L_t(\mathbf{p}). \tag{21}$$

Since the above is satisfied with equality only in the limit when $\delta \to 0$, the quantity $L_t(\mathbf{p} + \delta u\mathbf{x})u/\delta$ represents an approximation (possibly biased) of the directional derivative of $L_t(\mathbf{p})$ with respect to $\mathbf{x}$.

Now, in order to build a gradient estimate, the idea is to uniformly sample the objective function along a vector $\mathbf{u}(t)$ drawn from the S-dimensional Euclidean sphere of of radius δ. Extending the above to the space of dimension S, the estimator becomes:

$$\tilde{\mathbf{v}}(t) = \frac{S}{\delta} L_t(\tilde{\mathbf{p}}(t))\mathbf{u}(t), \tag{22}$$

where $\tilde{\mathbf{p}}(t) = \mathbf{p}(t) + \delta\mathbf{u}(t)$ and $\mathbf{u}(t)$ is uniformly taken over the unit Euclidean sphere: $\{\mathbf{u} \in \mathbb{R}^S |\ \|\mathbf{u}(t)\|^2 = 1\}$ [6]. More details are provided in Appendix E.

In [6, 7, 50], this estimator is proposed without accounting for the fact that the random sample point $\tilde{\mathbf{p}}(t) = \mathbf{p}(t) + \delta\mathbf{u}(t)$ can fall outside of the feasible set. In our power allocation problem, using the same procedure would imply that the transmit power vector $\tilde{\mathbf{p}}(t)$ is allowed to go outside $\mathcal{P}$. However, our power constraints are physical ones: transmit power positivity, maximum available power budget, which means that any violations are prohibited.

One of the major contributions of this work is to introduce a novel learning algorithm that exploits the gradient estimation above, while guaranteeing that the transmit powers always lie in the feasible set. For this, we define a modified and shrunk feasible set $\mathcal{P}_\delta$ such that, for any $\mathbf{p}_\delta(t) \in \mathcal{P}_\delta$, we have $\mathbf{p}_\delta(t) + \delta\mathbf{u}(t) \in \mathcal{P}$:

$$\mathcal{P}_\delta = \left\{ \mathbf{p}_\delta \in \mathbb{R}^S \;\middle|\; p_\delta^s \geq \delta, \sum_{s=1}^{S} p_\delta^s \leq P_{\max} - \sqrt{S}\delta \right\}. \tag{23}$$

Having defined this new feasible set, the suitable exponential map that guarantees that $\mathbf{p}_\delta(t) + \delta\mathbf{u}(t)$ always lies in this set is

$$p_\delta^s(t) \triangleq \delta + P_{\max}\,(1 - C_\delta)\,\frac{\exp(y^s(t))}{1 + \sum_{i=1}^{S} \exp(y^i(t))}, \tag{EXPδ}$$

where $C_\delta = \frac{\delta}{P_{\max}}(S + \sqrt{S})$. Using (EXP$\delta$), we introduce a novel exponential mapping: $\mathbf{Q}_\delta(\mathbf{y}(t)) \triangleq (p_\delta^1(t), \ldots, p_\delta^S(t))$. From the definition of $\mathcal{P}_\delta$ and (EXPδ), we can deduce the following conditions restricting the choice of the δ parameter

$$0 < \delta \leq \frac{P_{\max}}{S + \sqrt{S}} \leq \frac{P_{\max}}{\sqrt{S}}. \tag{24}$$

Summing up all ingredients, our novel algorithm can be summarized by the following three steps:

$$\begin{aligned} \tilde{\mathbf{v}}(t) &= \frac{S}{\delta} L_t(\tilde{\mathbf{p}}(t))\mathbf{u}(t), \\ \mathbf{y}(t+1) &= \mathbf{y}(t) - \mu\,\tilde{\mathbf{v}}(t) \\ \mathbf{p}_\delta(t) &= \mathbf{Q}_\delta(\mathbf{y}(t)), \end{aligned} \tag{OXL$_0$}$$

where $\tilde{\mathbf{v}}(t)$ represents the biased estimate of the gradient. For implementation details, see OXL$_0$ algorithm below. Although OXL$_0$ requires an additional step (i.e., the computation of the gradient estimation $\tilde{\mathbf{v}}(t)$) compared with OXL, the complexity of each iteration remains linear in the problem dimensionality S.

OXL$_0$ algorithm: Online Exponential Learning Algorithm with zeroth-order Feedback

Parameters: $\mu > 0; 0 < \delta \leq P_{max}/(S + \sqrt{S})$.
Initialization: $\mathbf{y}(0) \leftarrow 0; t \leftarrow 0$.
Repeat
- ▹ Pre-transmission phase:
- Update $\mathbf{p}_\delta(t) \leftarrow \mathbf{Q}_\delta(\mathbf{y}(t))$ defined in (EXPδ)
- Draw a random $\mathbf{u}(t)$ uniformly from the unit-sphere
- ▹ Transmit at $\tilde{\mathbf{p}}(t) \leftarrow \mathbf{p}_\delta(t) + \delta\mathbf{u}(t)$
- ▹ Post-transmission phase: receive scalar feedback $L_t(\tilde{\mathbf{p}}(t))$
- Compute the gradient estimation $\tilde{\mathbf{v}}(t) = \frac{S}{\delta} L_t(\tilde{\mathbf{p}}(t))\,\mathbf{u}(t)$
- Update scores $\mathbf{y}(t+1) \leftarrow \mathbf{y}(t) - \mu(t)\,\tilde{\mathbf{v}}(t)$
- $t \leftarrow t + 1$

until transmission ends

In Appendix E, we prove that the regret can be bounded as follows:

Theorem 4. *If the OXL$_0$ algorithm is run with constant parameters δ and μ then the average regret is bounded by:*

$$\begin{aligned} \mathbb{E}[Reg(T)] \leq \frac{P_{\max}\log(1+S)}{2\mu} + \mu T S^2 \left(\frac{B}{\delta} + K\right)^2 \\ + KT\delta\,\left(3 + P_{max}\left(S + 2\sqrt{S}\right)\right). \end{aligned} \tag{25}$$

where K is the Lipschitz constant and B the maximum value of the objective function $L_t(\cdot)$.

Tuning the parameters μ and δ: The step-size μ impacts the sensitivity of the algorithm to variations in the power policy. When μ is large, a small variation in the score $\mathbf{y}(t)$ results in a large variation in the power allocation. These large variations, can create oscillations in the power allocation policy $\mathbf{p}_\delta(t)$ and the time required to reach no regret increases as a result. At the opposite, a small μ leads to smaller variations in the power allocation, which also imply a long time for the regret to reach zero. Hence, there is a compromise and μ has to be carefully tuned to minimize the time to reach the no regret state.

The parameter δ represents the sampling radius of the device around the power policy $\mathbf{p}_\delta(t)$. When tuning δ, there is also a trade-off to be made between the precision of the gradient estimate and its variance. By reducing δ, the device reduces the distance to $\mathbf{p}_\delta(t)$ and the estimator gains in precision. But since the device only has access to one value of this estimate, reducing δ also increases the variability of the estimator (21).

The bound (25) can be further optimized, but because of the additional constraint $\delta \leq P_{max}/S + \sqrt{S}$, the resulting optimal bound will not be in closed-form. Having a slightly sub-optimal but closed-form expression will prove to be very useful in the sequel (when the time horizon T is unknown). For this, we choose $\delta^* = \frac{P_{max}}{(S+\sqrt{S})T^{1/4}}$ that always

meets the constraint and that decays optimally with respect to T. Then, we optimize the resulting bound in (25) only w.r.t to μ. The optimal μ^* is obtained by setting to zero the first-order derivative of the bound with respect to μ:

$$\mu^* = \sqrt{\frac{P_{max}\log(1+S)}{2T}}\left[S\left(\frac{B}{\delta^*}+K\right)\right]^{-1}. \quad (26)$$

Then, introducing δ^* and μ^* in (25) yields the bound

$$\mathbb{E}[Reg(T)] \leq U_1\, T^{3/4} + U_2\, T^{1/2}, \quad (27)$$

where

$$U_1 = S\,B\left(S+\sqrt{S}\right)\sqrt{\frac{2\log(1+S)}{P_{max}}} + K\left(3+P_{max}\left(S+2\sqrt{S}\right)\right)\frac{P_{max}}{S+\sqrt{S}}, \quad (28)$$

$$U_2 = \sqrt{2P_{max}\log(1+S)}\,S\,K. \quad (29)$$

Notice that the optimal bound w.r.t. δ and μ is also a function $\mathcal{O}(T^{3/4})$ and, hence, our particular choice of $\delta^*(T)$ above does not incur a large loss in terms of regret minimization rate and has the advantage of providing a closed-form expression of the upper bound.

Corollary 3. *If the OXL$_0$ algorithm is run for a known transmission horizon T and with the parameters δ^* and μ^* in* (26)*, then it leads to no regret and the average regret $\mathbb{E}[Reg(T)]/T$ vanishes as $\mathcal{O}(T^{-1/4})$.*

As in the previous sections, this result relies on the fact that the devices know their transmission horizon T in advance. To remove this requirement, the device can use the doubling trick or a time varying step-size. Since the time varying step-size generally involves a loss in the decay rate of the regret (see Sec. IV), we next investigate whether the information required by the doubling trick is readily available or not.

To do so, we have to determine specific values for the constants B and K in (26). A short calculation shows that they depend only on readily available system parameters:

$$B = S\,P_{\max} + \lambda R_{min}, \qquad K = 1 + 2\lambda R_{min}. \quad (30)$$

From (30) and (26) we conclude that the device is able to compute the parameters μ^* and δ^*. This implies that, if the time horizon T is not known in advance, the device can use the doubling trick described in Sec. III.

Proposition 2. *Assuming that the OXL$_0$ algorithm is run when T is unknown by using the doubling trick with the parameters μ_m^* and δ_m^* chosen as above in each window of size T_m, then the expected regret is bounded by*

$$\mathbb{E}[Reg(T)] \leq \frac{2\sqrt{2}}{2^{3/4}-1}\,U_1\,T^{3/4} + \frac{2}{\sqrt{2}-1}U_2\,T^{1/2}, \quad (31)$$

with U_1 and U_2 defined in (27)*. This means that the OXL$_0$ algorithm leads to no regret and the average regret $\mathbb{E}[Reg(T)]/T$ decreases at $\mathcal{O}(T^{-1/4})$.*

The proof follows similarly to the proof of Proposition 1 and is omitted. Importantly, reducing the available feedback results in a slower decay rate of the regret; the average regret vanishes as $\mathcal{O}(T^{-1/4})$ with zeroth-order feedback, whereas it vanishes as $\mathcal{O}(\log(T)T^{-1/2})$ with imperfect gradient feedback and as $\mathcal{O}(T^{-1/2})$ with perfect gradient feedback. Nevertheless, even under extremely limited feedback information - requiring a single sample of the objective function instead of its gradient - our proposed learning procedure (OXL$_0$ algorithm) achieves no-regret, irrespective of the evolution the network over time and despite the fact that its governing dynamics are unknown at the device end.

VI. Numerical experiments

Our goal in this section is to illustrate the performance guarantees of our learning algorithms in highly dynamic networks with realistic fading channel conditions, and with various degrees of (strictly causal, no look-ahead) feedback available at the device end, ranging from perfect gradient information to the bare-bones observation of the achieved loss. We start by comparing the OXL algorithm (full information) to classical approaches based on water-filling [21–23], suitably adapted to the setting at hand.

At each device, the benchmark water-filling is implemented so that the overall power consumption is minimized under the minimum rate constraint R_{min}. If the obtained solution does not meet the maximum power constraint, two possibilities are considered: a) the device remains silent - the *energy-driven* solution labeled WF0; b) the device transmits anyway by splitting the overall power budget uniformly over the S subcarriers - the *rate-driven* solution labeled WFP$_{\max}$.

We consider at first a simple setting composed of a pair of transmit-receive devices $N = M = 1$ communicating over four subcarriers ($S = 4$). The system parameters are: $\sigma^2 = 0.1$, $P_{max} = 1.5$ W, $R_{min} = 3$ bps/Hz and $\lambda = 1$; the channel gains are generated randomly as follows: $h^s(t+1) = \alpha h^s(t) + (1-\alpha)\varepsilon^s(t)$ with i.i.d. variables $\varepsilon^s(t) \sim \mathcal{N}(0,\sigma_\varepsilon^2)$ and $\sigma_\varepsilon^2 = 10$. This particular model allows us to control the temporal correlation of the channels via the parameter $\alpha \in [0,1]$ in between the extremes: the static channel case for $\alpha = 1$ (completely predictable); and the i.i.d. Rayleigh-fading case for $\alpha = 0$ (unpredictable).

For a fair comparison, we assume that the transmitting device only has access to a strictly causal feedback at each time instant. Fig. 3 illustrates the performance in terms of the relative outage defined as:

$$\text{Out} = [1 - R(\mathbf{p})/R_{\min}]^+, \quad (32)$$

of WF0 (Fig. 3(a)) and OXL algorithm (Fig. 3(b)). The performance is averaged over 100 realizations of the channel gains and for three different values of the time-correlation factor $\alpha \in \{0.2, 0.5, 0.8\}$. We remark that WF0 exhibits a high sensitivity to the temporal correlation of the channels:

Number of users	$M = 10$
Number of subcarriers	$S = 4$
Central frequency	$f_c = 2$ GHz
Bandwidth	10 MHz
Maximum power	$[0.5, 2]$ W
Minimum rate	$[0.5, 3]$ bps/Hz
λ	$[0.5, 10]$

Table I: Network parameters.

lower α (less predictable channel conditions), the worse the performance of WF-based algorithms. This is explained by the fact that water-filling algorithms perform well assuming that the SINR in each carrier is perfectly known ahead of the transmission (in the static channel case). Hence, the absence of look-ahead (non-causal) information negatively impacts the performance of classical water-filling algorithms. By contrast, the OXL algorithm consistently outperforms WF0 in terms of relative outage and is significantly more robust w.r.t. the channel dynamics. We find this feature of OXL to be particularly promising and appealing for applications to IoT networks, where the system changes constantly (and unpredictably), rendering conventional WF-based techniques obsolete.

The simple channel model above allowed us to highlight the impact of the channel dynamics and of having strictly causal feedback information on the system parameters. To validate the performance of OXL in more realistic environments, we consider in what follows a network composed of multiple interfering devices, in which the different channels are generated according to the commonly used COST-HATA model [51] that includes pathloss, fast fading and shadowing effects [47]. The speed of the devices is chosen arbitrarily between 0 km/h and 130 km/h so as to account for a wide spectrum of wireless mobile devices (smartphones, wearable, pedestrian, vehicle etc.). The minimum rate requirement R_{min}, the available power budget $P_{\max}$, and the rate vs. power tradeoff parameter λ also differ from one device to another.

Fig. 4 illustrates the comparison in terms of the rate vs. power consumption between OXL algorithm and both water-filling algorithms (Fig. 4(b) in the multiple device setting composed of $M = 10$ interfering devices over $S = 4$ subcarriers and communicating to the same receiver $N = 1$. The plotted curves are averaged over 100 realizations of the COST-HATA channel gains. We assume that all devices employ the same algorithms but with different parameters (for the OXL algorithm case).

We first note that classical water-filling is more rigid in terms of the rate vs. power tradeoff: either the device remains silent (WF0) or transmits with full power whenever its minimum rate constraint is incompatible with its power budget ($\text{WFP}_{\max}$). The parameter λ allows the device using OXL algorithm to smoothly tune its rate vs. power operating point depending on the target application. By increasing λ, the power consumption increases but the relative outage decreases. When all devices employ a rate-driven water-filling $\text{WFP}_{\max}$ a cascading effect emerges due to the fact that all devices are forced to transmit at full power ($p^s = P_{\max}/S$), which generates high network interference and, hence, has a deleterious effect on the algorithm's performance. We then see that both water-filling algorithms perform equally poorly in terms of relative outage when compared with the OXL algorithm (caused by their lack of robustness to strictly causal feedback information).

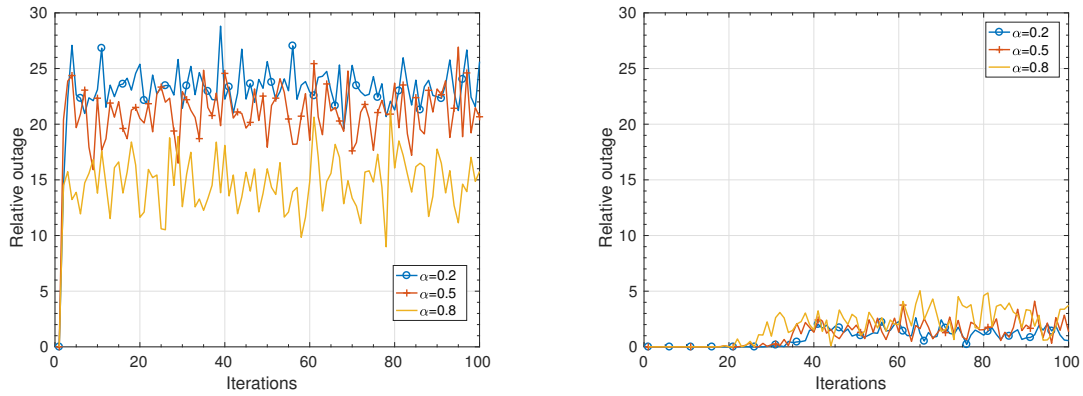
The next two goals of this section are: a) to validate our theoretical results in terms of regret, which evaluates both how close and how fast the proposed online algorithms reach the optimal fixed target state; and b) to investigate the effects of reducing the feedback information on the regret decay rate of the proposed methods.

Fig. 5 illustrates the vanishing regret of both our proposed algorithms, OXL (with perfect and imperfect gradient feedback) and OXL_0 (with a scalar feedback). Moreover, it also illustrates the impact of having a scarce or imperfect feedback and the impact of the problem dimensionality S. Fig. 5(a) confirms that having an imperfect gradient feedback does not influence significantly the regret decay rate, as anticipated by our theoretical results. However, this is no longer true when the only information available at the device end is a single scalar. The average regret of the OXL_0 algorithm decays slower compared with OXL algorithm (though the latter cannot be applied with zeroth-order feedback). Finally, Fig. 5(b) illustrates the average regret of OXL_0 algorithm for different values of the problem's dimensionality $S \in \{1, 2, 4\}$. In all cases, the average regret decays to zero; however if the number of available subcarriers increases, the variance of the estimator $\tilde{\mathbf{v}}(t)$ increases commensurately. Therefore the quality of the estimator decreases, which results in a reduced decay rate of the average regret.

VII. Conclusions

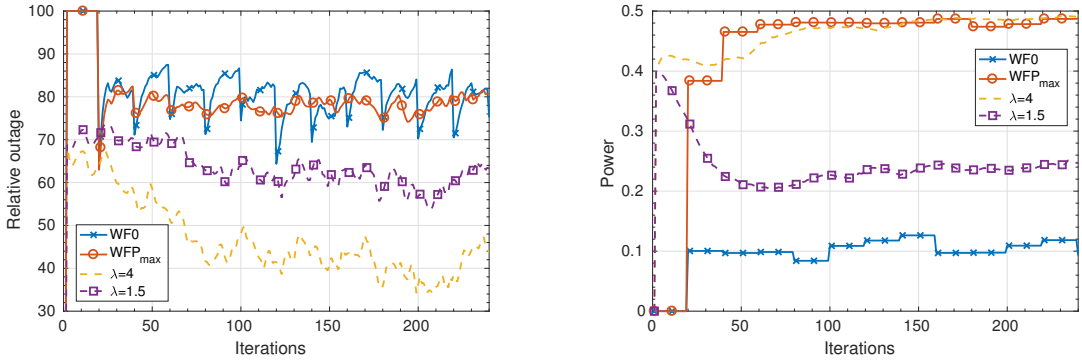
In this paper, we derived two adaptive algorithms (namely OXL and OXL_0) for solving power allocation problems in highly dynamic and unpredictable IoT networks based on online optimization tools and exponential learning. A key contribution lies in the fact that the proposed OXL_0 algorithm only requires the observation of a loss value at the device end. This algorithm is the first power allocation policy over multiple subcarriers which relies on a single scalar, as opposed to a vector worth of information containing the SINR values in all subcarrier required by classic water-filling algorithms.

Our simulations validate our theoretical expectations by showing that water-filling algorithms are highly sensitive to outdated feedback information and, hence, are not robust to rapid and unpredictable changes in the network. The proposed OXL algorithm outperforms classic water-filling algorithms in all investigated settings in which the network dynamics is not known at the device end. The impact of feedback scarcity is then assessed: the zeroth-order feedback algorithm is the slowest to reach no regret, followed



(a) Relative outage of WF0 (b) Relative outage of OXL algorithm

Figure 3: Performance comparison between WF0 and OXL algorithm in a time-varying setting, in which the transmitting device has access to strictly causal information. OXL algorithm outperforms WF0 in terms of relative outage irrespective from the channel dynamics. WF0 is negatively impacted by the outdated feedback information: the more unpredictable the channel gains, the higher the outage.



(a) Relative outage (b) Overall power

Figure 4: Performance comparison between OXL algorithm and water-filling for an arbitrary device. Water-filling algorithms are rigid in terms of the rate vs. power tradeoff, while OXL algorithm allows for a more smooth tuning via the parameter λ and for better performance both in terms of relative outage and power consumption. The rate-driven $\text{WFP}_{\max}$ exhibits a cascading effect in the network and forces all devices to transmit at full power resulting in poor performance. The energy-driven WF0 results in poor performance in terms of relative outage.

by the first algorithm in the imperfect gradient feedback case and then by the same algorithm in the perfect gradient case.

Appendix

A. First order feedback: known horizon T

For simplicity of presentation in the remaining appendices, we focus on the regret w.r.t. an arbitrary fixed policy $\mathbf{q} \in \mathcal{P}$ defined as $Reg_{\mathbf{q}}(T) \triangleq \sum_{t=1}^{T} L_t(\mathbf{p}(t)) - \sum_{t=1}^{T} L_t(\mathbf{q})$, and derive upper-bounds that are independent from $\mathbf{q}$ and, hence, also hold for the regret in (6) (or for its expectation).

The first step to prove Theorem 1 is to bound the regret based on the convexity of $L_t(\mathbf{q})$ as follows

$$Reg_{\mathbf{q}}(T) \leq \sum_{t=1}^{T} \langle \mathbf{v}(t) | \mathbf{p}(t) - \mathbf{q} \rangle \,, \tag{33}$$

where $\mathbf{q} \in \mathcal{P}$ is an arbitrarily chosen power allocation.

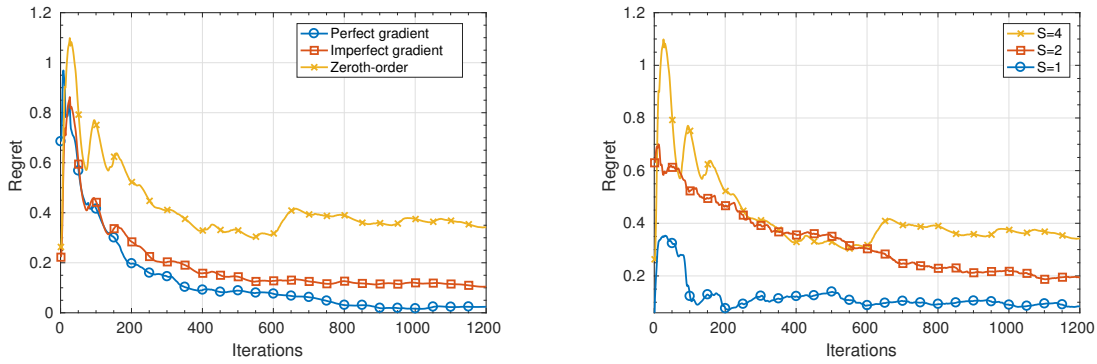
Using the fact that $\mathbf{y}(t+1) = \mathbf{y}(t) - \mu \mathbf{v}(t)$ and $\mathbf{y}(1) = 0$, we obtain

$$Reg_{\mathbf{q}}(T) \leq \sum_{t=1}^{T} \langle \mathbf{v}(t) | \mathbf{p}(t) \rangle + \frac{1}{\mu} \langle \mathbf{y}(T+1) | \mathbf{q} \rangle \,. \tag{34}$$

Then, we define a potential function $f^*(\mathbf{y}(t)) = P_{\max} \log(1 + \sum_{s=1}^{S} \exp(y_s(t)))$, which is used to show that the exponentiation step in (OXL) is equivalent to $\mathbf{p}(t) = \nabla f^*(\mathbf{y}(t))$. Also, the second order Taylor approximation of $f^*(\mathbf{y}(t))$ yields

$$f^*(\mathbf{y}(t+1)) \leq f^*(\mathbf{y}(t)) - \mu \langle \mathbf{v}(t) | \nabla f^*(\mathbf{y}(t)) \rangle + \frac{\mu^2}{2} P_{\max} \|\mathbf{v}(t)\|_2^2 . \tag{35}$$

Combining the above inequality with equation (34) and



(a) Regret of OXL and OXL_0 algorithms (b) Regret of OXL_0 algorithm

Figure 5: Impact of feedback amount and problem dimensionality. Both proposed algorithms, OXL and OXL_0, exhibit a vanishing regret, as anticipated by our theoretical results. The average regret of OXL_0 algorithm, relying only on the scalar value of the objective function, decays slower than the average regret of OXL algorithm with perfect or imperfect gradient feedback. Having to estimate the gradient of dimension S using the scalar value of the objective impacts the decay rate of the average regret of OXL_0 algorithm: the higher the problem dimensionality S, the slower the average regret.

given that $\mathbf{p}(t) = \nabla f^*(\mathbf{y}(t))$, we obtain:

$$Reg_{\mathbf{q}}(T) \leq \frac{1}{\mu}\left[f^*(0) - f^*(\mathbf{y}(T+1))\right] + \frac{\mu}{2}P_{\max}\sum_{t=1}^{T}\|\mathbf{v}(t)\|_2^2 + \frac{1}{\mu}\langle \mathbf{y}(T+1)|\mathbf{q}\rangle\,. \tag{36}$$

By using Fenchel's inequality [52] we get

$$f^*(\mathbf{y}) + f(\mathbf{q}) \geq \langle \mathbf{y}|\mathbf{q}\rangle\,,\ \forall\ \mathbf{y},\mathbf{q} \tag{37}$$

where $f(\mathbf{q})$ is the convex conjugate of $f^*(\mathbf{y})$ defined as $f(\mathbf{q}) = \sup_{\mathbf{y}\in\mathbb{R}}\langle \mathbf{y}|\mathbf{q}\rangle - f^*(\mathbf{y})$. We can then substitute $\langle \mathbf{y}(T+1)|\mathbf{q}\rangle - f^*(\mathbf{y}(T+1))$ by $f(\mathbf{q})$ in (36) and obtain

$$Reg_{\mathbf{q}}(T) \leq \frac{1}{\mu}\left[f(\mathbf{q}) + P_{\max}\log(1+S)\right] + \frac{\mu}{2}P_{\max}VT. \tag{38}$$

We can show that $f(\mathbf{q}) \leq 0$ for all $\mathbf{q} \in \mathcal{P}$ by using a variable change ($\mathbf{x} = \mathbf{q}/P_{\max}$) combined with Jensen's inequality for convex functions and the regret bound reduces to

$$Reg(T) \leq \frac{P_{\max}\log(1+S)}{\mu} + \frac{\mu}{2}P_{\max}TV \tag{39}$$

The optimal step-size is then obtained by minimizing the above bound.

B. First order feedback: unknown horizon T

OXL algorithm is run with the optimal step $\mu^*(T_m)$ defined in (9) in each window. Then, the regret in window m of size $T_m = 2^m$, denoted by $\widetilde{Reg}(T_m)$, can be bounded as in (10):

$$\widetilde{Reg}(T_m) \leq P_{\max}\sqrt{2T_mV\log(1+S)}. \tag{40}$$

For a time horizon T, the number of widows equals $\lceil\log_2(T)\rceil$, where $\lceil x\rceil$ is the ceiling function. The overall regret can be bounded by the sum of all windows' regrets:

$$Reg(T) \leq \sum_{m=0}^{\lceil\log_2(T)\rceil} P_{\max}\sqrt{2T_mV\log(1+S)}. \tag{41}$$

The result then follows by a geometric series argument.

C. Imperfect gradient feedback: known horizon T

From the convexity of the objective function, we can write

$$\mathbb{E}[Reg_{\mathbf{q}}(T)] \leq \mathbb{E}\left[\sum_{t=1}^{T}\langle \nabla L_t(\mathbf{p}(t))|\mathbf{p}(t)-\mathbf{q}\rangle\right]. \tag{42}$$

The idea is to link the above bound to the estimate $\tilde{\mathbf{v}}(t)$. By definition, we have that $\nabla L_t(\mathbf{p}(t)) = \mathbb{E}[\tilde{\mathbf{v}}(t)|\tilde{\mathbf{v}}(t-1),...,\tilde{\mathbf{v}}(1)]$. By the law of total expectation, the following equality holds

$$\mathbb{E}\left[\sum_{t=1}^{T}\langle \nabla L_t(\mathbf{p}(t))|\mathbf{p}(t)-\mathbf{q}\rangle\right] = \mathbb{E}\left[\sum_{t=1}^{T}\langle \tilde{\mathbf{v}}(t)|\mathbf{p}(t)-\mathbf{q}\rangle\right]. \tag{43}$$

The term inside the expectation on the RHS can be bounded as in (34) and, similarly to the proof of Therorem 1, we obtain

$$\mathbb{E}[Reg(T)] \leq \mathbb{E}\left[\frac{P_{\max}\log(1+S)}{\mu} + \frac{\mu}{2}P_{\max}\sum_{t=1}^{T}\|\tilde{\mathbf{v}}(t)\|_2^2\right]. \tag{44}$$

Given that $\mathbb{E}\left[\|\tilde{\mathbf{v}}(t)\|_2^2\right] \leq \tilde{V}$, the result follows.

D. Imperfect gradient feedback: unknown horizon T

To bound the regret assuming a variable step-size $\mu(t)$, we will consider the following weighted regret

$$WReg_{\mathbf{q}}(T) \triangleq \mathbb{E}\left[\sum_{t=1}^{T}\mu(t)\left(L_t(\mathbf{p}(t)) - L_t(\mathbf{q})\right)\right], \tag{45}$$

where $\mu(t)$ is the variable step-size in (17). Using a similar approach as in the proof of Theorem 3, we obtain

$$WReg_{\mathbf{q}}(T) \leq P_{\max}\log(1+S) + \frac{P_{\max}}{2}\tilde{V}\sum_{t=1}^{T}\mu^2(t). \quad (46)$$

To bound the regret, we use the summability criterion of Hardy [53], which allows us to compare weighted sums – here, $\mathbb{E}[Reg_{\mathbf{q}}(T)]$ and $WReg_{\mathbf{q}}(T)$. In particular, note that the step-size sequence $\mu(t) = \alpha t^{-1/2}$ satisfies the conditions $\mu(t) \geq \mu(t+1)$; and $\sum_{t=1}^{T}\mu(t)/\mu(T) = \mathcal{O}(T)$. Therefore, by Theorem 14 in [53], we obtain

$$\frac{\mathbb{E}[Reg_{\mathbf{q}}(T)]}{T} \sim \frac{WReg_{\mathbf{q}}(T)}{\sum_{t=1}^{T}\mu(t)} \leq \frac{P_{\max}}{\sqrt{T}}\left[\frac{\log(1+S)}{\alpha} + \frac{\alpha\tilde{V}(1+\log T)}{2}\right]. \quad (47)$$

E. Zeroth-order feedback: known horizon T

To prove Theorem 4, we introduce first some properties. Consider the following expectation of the objective function [7]

$$\tilde{L}_t(\mathbf{p}) \stackrel{\Delta}{=} \mathbb{E}_{\mathbf{u}\in\mathcal{B}}\left[L_t(\mathbf{p}+\delta\mathbf{u})\right], \quad (48)$$

where $\mathbf{u}$ is a random vector drawn uniformly on the unit Euclidean ball $\mathcal{B} = \left\{\mathbf{u}\in\mathbb{R}^s \,|\, \|\mathbf{u}\|^2 \leq 1\right\}$ and the expectation is taken over the randomness of $\mathbf{u}$. We can show that $\tilde{L}_t(\cdot)$ is a biased estimator of $L_t(\cdot)$ and

$$|L_t(\mathbf{p}) - \tilde{L}_t(\mathbf{p})| \leq K\delta, \quad \forall\mathbf{p}, \quad (49)$$

where K is the Lipschitz constant of the objective function. An important property of $\tilde{L}_t(\mathbf{p})$ is that its gradient relies on the values of the objective function as follows

$$\nabla\tilde{L}_t(\mathbf{p}) = \mathbb{E}_{\mathbf{u}\in\mathcal{S}}\left[\frac{S}{\delta}L_t(\mathbf{p}+\mathbf{u}\delta)\mathbf{u}\right], \quad (50)$$

where $\mathbf{u}$ is drawn for the unit Euclidean sphere $\mathcal{S} = \left\{\mathbf{u}\in\mathbb{R}^S \,|\, \|\mathbf{u}\|^2 = 1\right\}$.

Another useful property is that the new exponential mapping step in (EXPδ), which is adapted to the modified set $\mathcal{P}_\delta$, can be written equivalently as:

$$\mathbf{p}_\delta(t) = \arg\max_{\mathbf{q}\in\mathcal{P}_\delta}\left\{\langle\mathbf{y}(t)|\mathbf{q}\rangle - h(\mathbf{q})\right\},$$
$$h(\mathbf{q}) \triangleq \sum_{s=1}^{S}(q^s-\delta)\log(q^s-\delta) + \left(C - \sum_{s=1}^{S}q^s\right)\log\left(C - \sum_{s=1}^{S}q^s\right), \quad (51)$$

with $C = P_{\max} - \delta\sqrt{S}$.

The first step to prove Theorem 4 is to compare $L_t(\mathbf{p}_\delta(t)+\delta\mathbf{u})$, the incurred loss at time t, to $L_t(\mathbf{p}_\delta(t))$ by using that $L_t(\cdot)$ is a K-Lipschitz function:

$$\mathbb{E}[Reg_{\mathbf{q}}(T)] \leq \mathbb{E}\left[\sum_{t=1}^{T}L_t(\mathbf{p}_\delta(t)) - L_t(\mathbf{q})\right] + KT\delta\left(1 + P_{max}\tilde{S}\right),$$

with $\tilde{S} = S + 2\sqrt{S}$. The second step is to compare $L_t(\mathbf{p}_\delta(t))$ and $L_t(\mathbf{q})$ to $\tilde{L}_t(\mathbf{p}_\delta(t))$ and $\tilde{L}_t(\mathbf{q})$ respectively.

$$\mathbb{E}[Reg_{\mathbf{q}}(T)] \leq \mathbb{E}\left[\sum_{t=1}^{T}\tilde{L}_t(\mathbf{p}_\delta(t)) - \tilde{L}_t(\mathbf{q})\right] + KT\delta\left(3 + P_{max}\tilde{S}\right).$$

Since $\tilde{L}_t(\mathbf{p})$ is convex w.r.t. $\mathbf{p}$ we have:

$$\mathbb{E}\left[\sum_{t=1}^{T}\tilde{L}_t(\mathbf{p}_\delta(t)) - \tilde{L}_t(\mathbf{q})\right] \leq \mathbb{E}\left[\sum_{t=1}^{T}\langle\nabla\tilde{L}_t(\mathbf{p}_\delta(t))|\mathbf{p}_\delta(t)-\mathbf{q}\rangle\right].$$

We can write $\nabla\tilde{L}_t(\mathbf{p}_\delta(t)) = \mathbb{E}[\tilde{\mathbf{v}}(t)|\mathbf{u}(1),...,\mathbf{u}(t-1)]$, where $\tilde{\mathbf{v}}(t)$ is the estimation defined in (22) and where the expectation is taken over the randomness of $\mathbf{u}$. Using this property and the law of total expectation, the bound on the expected regret becomes:

$$\mathbb{E}\left[\sum_{t=1}^{T}\langle\nabla\tilde{L}_t(\mathbf{p}_\delta(t))|\mathbf{p}_\delta(t)-\mathbf{q}\rangle\right] \leq \mathbb{E}\left[\sum_{t=1}^{T}\langle\tilde{\mathbf{v}}(t)|\mathbf{p}_\delta(t)-\mathbf{q}\rangle\right].$$

By using (51), we can bound the sum $\sum_{t=1}^{T}\langle\tilde{\mathbf{v}}(t)|\mathbf{q}\rangle$ and obtain

$$\mathbb{E}\left[\sum_{t=1}^{T}\langle\tilde{\mathbf{v}}(t)|\mathbf{p}_\delta(t)-\mathbf{q}\rangle\right] \leq \mathbb{E}\left[\sum_{t=1}^{T}\langle\mathbf{v}(t)|\mathbf{p}_\delta(t+1)-\mathbf{p}_\delta(t)\rangle\right] + \frac{H}{2\mu},$$

where $H = \min_{\mathbf{p}\in\mathcal{P}_\delta}h(\mathbf{p})$. We use again (51) and the Cauchy-Schwartz inequality to bound the sum of $\langle\mathbf{v}(t)|\mathbf{p}_\delta(t+1)-\mathbf{p}_\delta(t)\rangle$ and, by combining all the above, we find

$$\mathbb{E}[Reg(T)] \leq \frac{H}{2\mu} + \mu TS^2\left(\frac{B}{\delta}+K\right)^2 + KT\delta\left(3 + P_{max}\left(S + 2\sqrt{S}\right)\right).$$

where $B = \max_{t,\mathbf{p}}L_t(\mathbf{p})$. Finally, Theorem 4 follows by finding that $H = P_{max}\log(1+S)$.

References

[1] A. Marcastel, E. V. Belmega, P. Mertikopoulos, and I. Fijalkow, "Online power allocation for opportunistic radio access in dynamic OFDM networks," in *Vehicular Technology Conference (VTC-Fall), 2016 IEEE 84th*, Sep. 2016, pp. 1–5.

[2] ——, "Online interference mitigation via learning in dynamic IoT environments," in *Globecom Workshop, Internet of Everything, 2016 IEEE*, Dec. 2016, pp. 1–5.

[3] A. I. Sulyman, S. M. Oteafy, and H. S. Hassanein, "Expanding the cellular-IoT umbrella: An architectural approach," *IEEE Trans. Wireless Commun.*, vol. 24, no. 3, pp. 66–71, Jun. 2017.

[4] M. Basharat, W. Ejaz, M. Naeem, A. M. Khattak, and A. Anpalagan, "A survey and taxonomy on nonorthogonal multiple-access schemes for 5G networks," *Transactions on Emerging Telecommunications Technologies*, vol. 29, no. 1, pp. 1–17, Jun. 2018.

[5] Huawei Technologies, *5G: A technology vision*. White paper, 2013. [Online]. Available: www.huawei.com/ilink/en/download/HW_314849

[6] S. Shalev-Shwartz, "Online learning and online convex optimization," *Foundations and Trends in Machine Learning*, vol. 4, no. 2, pp. 107–194, 2011.

[7] S. Bubeck, N. Cesa-Bianchi *et al.*, "Regret analysis of stochastic and nonstochastic multi-armed bandit problems," *Foundations and Trends in Machine Learning*, vol. 5, no. 1, pp. 1–122, 2012.

[8] I. Caragiannis, C. Kaklamanis, P. Kanellopoulos, M. Kyropoulou, B. Lucier, R. P. Leme, and E. Tardos, "Bounding the inefficiency of outcomes in generalized second price auctions," *Journal of Economic Theory*, vol. 156, pp. 343–388, Mar. 2015.

[9] N. Cesa-Bianchi, C. Gentile, and Y. Mansour, "Regret minimization for reserve prices in second-price auctions," *IEEE Trans. Inf. Theory*, vol. 61, no. 1, pp. 549–564, Jan. 2015.

[10] P. Viappiani and C. Boutilier, "Regret-based optimal recommendation sets in conversational recommender systems," in *Proc. 3rd ACM conference on Recommender Systems*, Oct. 2009, pp. 101–108.

[11] P. Mertikopoulos and E. V. Belmega, "Transmit without regrets: Online optimization in MIMO-OFDM cognitive radio systems," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 11, pp. 1987–1999, Dec. 2014.

[12] ——, "Learning to be green: Robust energy efficiency maximization in dynamic MIMO-OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 743–757, Apr. 2016.

[13] D. Miorandi, S. Sicari, F. De Pellegrini, and I. Chlamtac, "Internet of things: Vision, applications and research challenges," *Ad Hoc Networks*, vol. 10, no. 7, pp. 1497–1516, Sep. 2012.

[14] C. Goursaud and J.-M. Gorce, "Dedicated networks for IoT: PHY/-MAC state of the art and challenges," *EAI Endorsed Transactions on Internet of Things*, vol. 1, no. 1, pp. 1–11, Oct. 2015.

[15] Y. Chen, F. Han, Y.-H. Yang, H. Ma, Y. Han, C. Jiang, H.-Q. Lai, D. Claffey, Z. Safar, and K. R. Liu, "Time-reversal wireless paradigm for green internet of things: An overview," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 81–98, 2014.

[16] W. Li, M. Assaad, and P. Duhamel, "Distributed stochastic optimization in networks with low informational exchange," in *Communication, Control, and Computing (Allerton), 2017 55th Annual Allerton Conference on*. IEEE, 2017, pp. 1160–1167.

[17] D. J. Love, R. W. Heath, V. K. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, Oct. 2008.

[18] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Areas Commun.*, vol. 8, no. 5, pp. 742–758, Apr. 2014.

[19] W. Li, M. Assaad, G. Ayache, and M. Larranaga, "Matrix exponential learning for resource allocation with low informational exchange," in *IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2018, pp. 266–270.

[20] W. Li and M. Assaad, "Matrix exponential learning schemes with low informational exchange," *arXiv preprint arXiv:1802.06652*, 2018. [Online]. Available: https://arxiv.org/pdf/1802.06652v2.pdf

[21] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 145–152, Jan. 2004.

[22] J.-S. Pang, G. Scutari, F. Facchinei, and C. Wang, "Distributed power allocation with rate constraints in Gaussian parallel interference channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3471–3489, Jul. 2008.

[23] G. Scutari, D. P. Palomar, and S. Barbarossa, "The MIMO iterative waterfilling algorithm," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1917–1935, Jan. 2009.

[24] H. Safdar, N. Fisal, R. Ullah, W. Maqbool, F. Asraf, Z. Khalid, and A. Khan, "Resource allocation for uplink M2M communication: A game theory approach," in *Wireless Technology and Applications (ISWTA), 2013 IEEE Symp.*, Sep. 2013, pp. 48–52.

[25] M. S. Ali, H. Tabassum, and E. Hossain, "Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (NOMA) systems," *IEEE Access*, vol. 4, pp. 6325–6343, Aug. 2016.

[26] T. Zheng, Y. Qin, H. Zhang, and S. Kuo, "Adaptive power control for mutual interference avoidance in industrial Internet-of-Things," *China Communications*, vol. 13, no. Supplement 1, pp. 124–131, Sep. 2016.

[27] G. J. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithm and its convergence," *IEEE Trans. Veh. Technol.*, vol. 42, no. 4, pp. 641–646, 1993.

[28] R. Masmoudi, E. V. Belmega, I. Fijalkow, and N. Sellami, "A unifying view on energy-efficiency metrics in cognitive radio channels," in *Signal Processing Conference (EUSIPCO), 2014 Proc. 22nd European*, Sep. 2014, pp. 171–175.

[29] T. Holliday, N. Bambos, P. Glynn, and A. Goldsmith, "Distributed power control for time varying wireless networks: Optimality and convergence," in *Proc. of the annual ALLERTON Conference on Communication Control and Computing*, vol. 41, no. 2, 2003, pp. 1024–1033.

[30] R. Tajan, C. Poulliat, and I. Fijalkow, "Interference management for cognitive radio systems exploiting primary IR-HARQ: A constrained Markov decision process approach," in *Signals, Systems and Computers (ASILOMAR), 2012 Conference Record of the Forty Sixth Asilomar Conference on*, 2012, pp. 1818–1822.

[31] C. Isheden, Z. Chong, E. Jorswieck, and G. Fettweis, "Framework for link-level energy efficiency optimization with informed transmitter," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2946–2957, 2012.

[32] J. Wang, C. Jiang, Z. Han, Y. Ren, and L. Hanzo, "Network association strategies for an energy harvesting aided super-WiFi network relying on measured solar activity." *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3785–3797, 2016.

[33] T. Chen, A. Mokhtari, X. Wang, A. Ribeiro, and G. B. Giannakis, "Stochastic averaging for constrained optimization with application to online resource allocation," *IEEE Trans. Signal Process.*, vol. 65, no. 12, pp. 3078–3093, 2017.

[34] P. Mertikopoulos, E. V. Belmega, R. Negrel, and L. Sanguinetti, "Distributed stochastic optimization via matrix exponential learning," *IEEE Trans. Signal Process.*, vol. 65, no. 9, pp. 2277–2290, 2017.

[35] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, 2017.

[36] A. Anandkumar, N. Michael, A. K. Tang, and A. Swami, "Distributed algorithms for learning and cognitive medium access with logarithmic regret," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 4, pp. 731–745, Mar. 2011.

[37] M. Hashemi, A. Sabharwal, C. E. Koksal, and N. B. Shroff, "Efficient beam alignment in millimeter wave systems using contextual bandits," *arXiv preprint arXiv:1712.00702*, 2017.

[38] J. Wang, C. Jiang, H. Zhang, X. Zhang, V. C. Leung, and L. Hanzo, "Learning-aided network association for hybrid indoor LiFi-WiFi systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3561–3574, 2018.

[39] T. Chen, S. Barbarossa, X. Wang, G. B. Giannakis, and Z.-L. Zhang, "Learning and management for Internet-of-Things: Accounting for adaptivity and scalability," *arXiv preprint arXiv:1810.11613*, 2018.

[40] L. Mainetti, L. Patrono, and A. Vilei, "Evolution of wireless sensor networks towards the internet of things: A survey," in *Software, Telecommunications and Computer Networks (SoftCOM), IEEE 19th Intl. Conf. on*, 2011, pp. 1–6.

[41] G. M. Lee and N. Crespi, "The Internet of Things: Challenge for a new architecture from problems," in *IAB Interconnecting Smart Objects with the Internet Workshop*, Mar. 2011, pp. 1–2.

[42] T. Alpcan, T. Başar, R. Srikant, and E. Altman, "CDMA uplink power control as a noncooperative game," *Wireless Networks*, vol. 8, no. 6, pp. 659–670, Nov. 2002.

[43] E. Altman and L. Wynter, "Equilibrium, games, and pricing in transportation and telecommunication networks," *Networks and Spatial Economics*, vol. 4, no. 1, pp. 7–21, Mar. 2004.

[44] M. Chiang, P. Hande, T. Lan, C. W. Tan *et al.*, "Power control in wireless cellular networks," *Foundations and Trends in Networking*, vol. 2, no. 4, pp. 381–533, 2008.

[45] E. V. Belmega, P. Mertikopoulos, R. Negrel, and L. Sanguinetti, "Online convex optimization and no-regret learning: Algorithms, guarantees and applications," *arXiv preprint arXiv:1804.04529, submitted to IEEE Signal Processing Magazine*, 2018.

[46] M. Zinkevich, "Online convex programming and generalized infinitesimal gradient ascent," in *Proc. of the 20th International Conference on Machine Learning (ICML-03)*, Aug. 2003, pp. 928–936.

[47] G. Calcev, D. Chizhik, B. Göransson, S. Howard, H. Huang, A. Kogiantis, A. F. Molisch, A. L. Moustakas, D. Reed, and H. Xu, "A wideband spatial channel model for system-wide simulations," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 389–403, Mar. 2007.

[48] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge University Press, 2005.

[49] L. B. Klebanov, S. T. Rachev, and F. J. Fabozzi, *Robust and Non-Robust models in Statistics*. Nova Science Publishers, 2009.

[50] A. D. Flaxman, A. T. Kalai, and H. B. McMahan, "Online convex optimization in the bandit setting: gradient descent without a gradient," in *SODA'05: Proceedings of the 16th annual ACM-SIAM symposium on discrete algorithms*, Jan. 2005, pp. 385–394.

[51] G. F. Pedersen, *COST 231-Digital mobile radio towards future generation systems*. EU, 1999.

[52] R. T. Rockafellar, *Convex analysis*. Princeton University Press, 2015.

[53] G. H. Hardy, *Divergent Series*. Oxford University Press, 1949.

Alexandre Marcastel (S'16) received his engineer degree from ENSEA, Cergy-Pontoise, France in 2014, the M.Sc. and Ph.D. degrees both from the University of Cergy-Pontoise, France in 2015 and 2019, respectively. During his Ph.D. in ETIS laboratory (ETIS, UMR 8051, University Paris Seine, University Cergy-Pontoise, ENSEA, CNRS), he worked on applications of online optimization to dynamic and unpredictable wireless (IoT) networks. He is currently a teaching assistant in ENSEA, Cergy-Pontoise, France and his main research interests lie in resource allocation problems for dynamic wireless networks.

E. Veronica Belmega (S'08–M'10) received the M.Sc. (engineer diploma) degree from the University Politehnica of Bucharest, Bucharest, Romania, in 2007, and the M.Sc. and Ph.D. degrees both from the University Paris-Sud 11, Orsay, France, in 2007 and 2010, respectively. From 2010 to 2011, she was a Postdoctoral Researcher in a joint project between Princeton University, N.J., USA and the Alcatel-Lucent Chair on Flexible Radio in Supélec, France. She is currently an Associate Professor with ETIS/ENSEA–Université de Cergy-Pontoise–CNRS, Cergy-Pontoise, France. She was one of the ten recipients of the L'Oréal–UNESCO–French Academy of Science Fellowship: "For young women doctoral candidates in science" in 2009. Since 2018, she is the recipient of the Doctoral Supervision and Research Bonus (PEDR) by the French National Council of Universities (CNU 61). She serves on the editorial board of the Transactions on Emerging Telecommunications Technologies (ETT) and has been distinguished among the Top 11 Editors, for outstanding contributions to ETT during 2016-2017.

Panayotis Mertikopoulos (M' 11) received the Ptychion degree in physics (summa cum laude) from the University of Athens in 2003, his M.Sc. and M.Phil. degrees in mathematics from Brown University in 2005 and 2006 (both summa cum laude), and his Ph.D. degree from the University of Athens in 2010. During 2010–2011, he was a post-doctoral researcher at the École Polytechnique, Paris, France. Since 2011, he has been a CNRS Researcher at the Laboratoire d'Informatique de Grenoble, Grenoble, France. P. Mertikopoulos was an Embeirikeion Foundation Fellow between 2003 and 2006, and received the best paper award in NetGCoop '12. He is serving on the editorial board and program committees of several journals and conferences on learning and optimization (such as NIPS and ICML). His main research interests lie in learning, optimization, game theory, and their applications to networks and machine learning systems.

Inbar Fijalkow (M'96–SM'10) received her engineering and Ph.D. degrees from TelecomParis, Paris, France, in 1990 and 1993, respectively. In 1993–1994, she was a postdoctoral research fellow at Cornell University, NY, USA, supported by a French Lavoisier Fellowship. Since 1994, she is a member of ETIS research unit, UMR 8051, Université Paris Seine, Université Cergy-Pontoise, ENSEA, CNRS, while teaching at ENSEA in Cergy, France. From 2000 to 2004, she was in charge of the master research program in intelligent and communicating systems (SIC). She was the vice-head of the French research group in image and signal processing, GdR ISIS, from 2002 to 2004. She was the dean of the ETIS laboratory from 2004 to 2013. In 2015-2016, she was a visiting researcher at UC Irvine (CA, USA) funded by the French CNRS. She is currently a full Professor at ENSEA at the "classe exceptionnelle" level. Her research interests are in signal processing for digital communications, including iterative processing, optimization, estimation theory, signal processing for dirty-RF. She is (co-)author of over 180 publications. Dr. Fijalkow is a past IEEE Transactions on Signal Processing associate editor. She is a Senior IEEE member and the member of several technical committees. She is a mentor in several associations for women in science and engineering.